



# Analyse de survie en présence d'hétérogénéité entre sujets dans les essais thérapeutiques

Elsa Cecilia-Joseph

## ► To cite this version:

Elsa Cecilia-Joseph. Analyse de survie en présence d'hétérogénéité entre sujets dans les essais thérapeutiques. Santé publique et épidémiologie. Université Paris-Saclay, 2015. Français. NNT : 2015SACLS206 . tel-01274076

**HAL Id: tel-01274076**

**<https://theses.hal.science/tel-01274076>**

Submitted on 15 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2015SACLS206

THESE DE DOCTORAT  
DE  
L'UNIVERSITE PARIS-SACLAY  
PREPAREE A  
L'UNIVERSITE PARIS-SUD

ÉCOLE DOCTORALE N°570  
SANTÉ PUBLIQUE  
Spécialité de doctorat : Biostatistiques

Par

**Melle CECILIA-JOSEPH Elsa**

Analyse de Survie en présence d'hétérogénéité entre sujets  
dans les essais thérapeutiques.

**Thèse présentée et soutenue au Kremlin-Bicêtre, le 7 décembre 2015 :**

**Composition du Jury :**

M. Denis HEMON, Professeur, Université Paris-Sud, Président du Jury  
M. John O'QUIGLEY, Professeur, Université Pierre et Marie Curie, Rapporteur  
M. Raphaël PORCHER, Professeur, Université Paris Diderot, Rapporteur  
Mme. Virginie RONDEAU, Directrice de Recherche, Université Bordeaux 2, Examinatrice  
M. Bertran AUVERT, Professeur, Université Versailles St Quentin, Directeur de thèse  
M. Thierry MOREAU, Directeur de Recherche, Université Paris-Sud, Directeur de thèse



# Remerciement

Dans un premier temps, je tiens à remercier chaleureusement mes directeurs de thèse, Thierry Moreau et Bertran Auvert de m'avoir fait confiance en m'offrant la possibilité de travailler sur ce sujet. Un grand merci à Thierry qui avant même les débuts officiels de cette thèse n'a pas douté de moi et dont les encouragements n'ont pas cessé tout le long de ces quatre années, avec toute la bienveillance qui le caractérise. Je suis d'ailleurs honorée d'être sa dernière thésarde avant une retraite que je lui souhaite douce et agréable. Un grand merci également à Bertran pour sa sympathie et tous ses précieux conseils qui m'ont permis à chaque fois d'avoir une meilleure vision de la réalité de terrain, une notion essentielle pour devenir un bon biostatisticien.

Je souhaite également remercier tous les membres du jury d'avoir accepté d'évaluer mon travail. Merci à Denis Hémon de me faire l'honneur de présider ce jury. Merci à mes rapporteurs John O'Quigley et Raphaël Porcher, de prendre le temps de lire ce travail et de me faire bénéficier de leur précieux conseils. Merci à Virginie Rondeau d'avoir bien voulu évaluer mon travail.

Je remercie vivement l'ensemble de l'équipe Biostatistique et Pharmaco-épidémiologie de l'Unité 1181, qui était alors l'équipe Biostatistique de l'Unité 1018 quand elle m'a si chaleureusement accueillie dès mon stage de fin de Master. Je remercie également toute l'équipe d'épidémiologie respiratoire et environnementale pour sa bonne humeur qui a laissé un grand vide derrière elle. Je remercie aussi Phillippe Broët pour son aide précieuse qui a contribué grandement à atteindre ma première publication.

Je remercie l'Université Paris-Sud pour avoir financé mes trois premières années de thèse, ainsi que l'Agence Nationale de Recherche sur le Sida qui a financé ma dernière année, me donnant ainsi l'opportunité de finaliser mes travaux.

Un merci particulier aux doctorants, post-doctorants... avec qui j'ai eu le plaisir de partager mes repas du midi et quelques mijitos du soir, et qui ont rendu ces années si agréables : Dorota, Helena, Juliette, Margarita, Mohammed, Fanny, Valérie, Orianne, Marta, Margaux, Annabelle, Emilie, Zhen, Cathy, Sabrina, Sylwester, Jonathan, Yves...

Un grand merci à mes amis qui me soutiennent bien au delà de la thèse. Merci de m'avoir particulièrement aidé pour les moments de détente : Mailly, Morgan, Gwenaëlle, Soraya, Nathalie, Yannis, Miguel, Joël...

Je tiens à remercier également Paul-Emile Maingé, l'un de mes professeurs à la faculté des Antilles et de la Guyane, qui a été la première personne à me motiver pour me lancer dans ce long périple qu'est la thèse, et qui n'a cessé de suivre son évolution jusqu'à son aboutissement.

Merci surtout ma famille pour son aide et son soutien sans faille. Merci Maman de ne pas hésiter à faire 8000km si le besoin peut s'en faire sentir. Merci Papa pour tes encouragements précieux. Merci Kelly pour ta bienveillance digne d'une grande « petite » sœur. Merci Yannick pour ta positive attitude contagieuse. Et enfin merci à Jérémy qui m'a supportée, encouragée, même quand ce n'était pas toujours facile, et qui m'a permis de finir ces quatre années avec bien plus qu'une thèse. Merci pour le « petit » cadeau qu'il m'a offert, dont les grands yeux et le sourire m'encourageaient aussi tous les jours depuis plus d'un an maintenant... Donc merci à toi aussi Jeylan.

# Résumé

Au cours des études de survie, certains facteurs ayant un rôle pronostique peuvent être inobservés ou indisponibles. Dans le cadre des essais cliniques randomisés où deux groupes de traitement sont comparés par un modèle de Cox, « l'oubli » de tels facteurs dans le modèle entraîne une sous-estimation en valeur absolue de l'effet du traitement. L'objectif de la thèse est de préciser le rôle de différents déterminants de ce biais et de suggérer l'utilisation de méthodes pouvant le réduire, avec un intérêt particulier pour les essais de prévention à l'infection VIH où de tels facteurs peuvent exister.

L'effet des principaux facteurs pouvant influencer le biais est précisé dans une première partie de la thèse. Si certains facteurs sont connus de longue date, d'autres, comme la durée de l'essai n'ont, à notre connaissance, pas été étudiés. L'approche utilisée repose sur l'utilisation d'un « vrai » modèle de type à risques proportionnels. Dans ce modèle, l'effet des covariables « oubliées » est résumé par l'introduction dans l'expression du risque instantané d'un terme aléatoire de « fragilité » propre à chaque sujet. Le biais cherché est calculé comme la limite asymptotique, sous le modèle marginal correspondant au « vrai » modèle, du rapport des risques instantanés déduit du modèle de Cox n'incluant pas la fragilité. Les résultats montrent une nette augmentation du biais, en valeur absolue, avec la durée de l'essai. Cette augmentation est particulièrement marquée pour des distributions de fragilité continues comme celles pouvant être rencontrées en pratique, par rapport à des fragilités binaires. Par ailleurs, les résultats

antérieurs de la littérature sont confirmés et précisés.

Dans un second temps, les conséquences d'une variation de la fragilité au cours du temps sur le biais sont recherchées. Plus précisément, la situation envisagée est celle qui est rencontrée dans les essais de prévention contre le VIH effectués auprès de populations « instables » vis-à-vis du risque, comme les jeunes ou les prostituées en Afrique sub-saharienne. Ces populations montrent une hétérogénéité comportementale liée aux changements de partenaires sexuels dont le statut VIH est par ailleurs inconnu. Il s'agit d'évaluer le biais en présence d'une telle fragilité « intermittente » au moyen de simulations reflétant les situations réelles. Les résultats montrent que le biais dû à l'omission de la fragilité dans l'analyse, bien qu'inférieur au biais obtenu dans le cas d'une fragilité constante au cours du temps, peut demeurer non négligeable.

Enfin, dans le cas d'une fragilité supposée constante au cours du temps, l'intérêt de l'utilisation d'un essai en « cross-over » est recherché. Dans un tel essai, les sujets exempts d'évènement après un temps de suivi donné changent de groupe de traitement. Dans le cadre des essais de prévention VIH, Auvert & al [2011] ont montré en particulier une diminution du biais avec un schéma en cross-over comparativement à un schéma parallèle classique, en utilisant une fragilité catégorielle dans une étude de simulation. Buyze & Goetghebeur [2011] ont également montré les avantages du cross-over, en particulier concernant l'efficacité relative d'un test de comparaison des deux groupes, en utilisant une fragilité de distribution gamma ou log-normale. Ces résultats sont précisés en calculant formellement le biais asymptotique dans l'estimation du risque relatif pour les différentes distributions possibles de la fragilité omise, continue ou catégorielle. Les résultats obtenus sont nettement en faveur du cross-over, avec une diminution du biais entre 60% et 90% et une amélioration de l'efficacité du test. Le temps de changement de groupe optimal est également recherché ainsi que les facteurs influant sur sa valeur. Il apparaît que celui-ci dépend essentiellement de la durée de l'essai et de la valeur de

l'effet traitement, et est très peu affecté par la distribution de la fragilité.

Mots clés : essais thérapeutiques, cross-over, modèle de Cox, fragilité, hétérogénéité, covariables oubliées, essais de prévention VIH.





# Abstract

In survival analysis, some prognostic factors can be unobserved or unavailable. In randomized clinical trials framework where two treatment groups are compared in the Cox model setting, the omission of such factors in the model leads to an underestimation in absolute value of the treatment effect. The aim of the project is to better understand the determinants of this bias and to suggest the use of methods that could reduce it, with a particular interest in HIV prevention trials where such factors are likely to exist.

In a first step, the role of the main determinants of the bias is highlighted. While some of them have long been identified, others like trial duration have never been considered to our knowledge. The bias was calculated as the asymptotic limit of the maximum likelihood estimator of the treatment effect when the analysis is done following a proportional hazard model which no takes into account the frailty. The results show a clear increase of the bias in absolute value with the trial duration. This increase is especially marked with continuous frailty distributions, such as those which can be encountered in practice, compared to binary frailties. Also, some previous results have been confirmed.

In a second step, a frailty depending on time is considered, as it can be encountered in HIV clinical trials including “unstable” population about infection risk, as prostitutes or young people in sub-Saharan Africa. These populations present a behavioral hetero-

geneity linked to the change of partner whose the HIV status is unknown. The bias is estimated using a transient frailty with a simulation study reflecting real-life situations. The results show that when the frailty is regenerated during the follow-up, the bias caused by its omission, although lower than the bias obtained with a time-independent frailty, stays significant and has to be considered.

Finally, with a “stable” population whose the frailty can be supposed constant over time, the interest of the use of the “cross-over” is searched. In such trial, the subjects which have not presented the event after a given time of follow-up, change of treatment group. In HIV prevention trials framework, Auvert & al. (2011) have particularly shown a decrease of the bias with a cross-over design comparatively to a parallel design, using a categorized frailty in a simulation study. Buyze & Goetghebeur (2011) have also shown the advantages of the cross-over, particularly about the relative efficiency of the comparison test of the two treatment groups, using gamma or log-normal frailty distributions. These results are specified calculating explicitly the asymptotic bias of the hazard ratio estimate for different possible distributions of the omitted frailty, continuous or categorized. The obtained results are clearly in favor of the cross-over, with a decrease of the bias between 60% et 90% and a significant improvement of the efficiency of the comparison test. The optimal switch time and its prognostic factors are searched. It appears that it essentially depends on the trial duration and the value of the treatment effect, and is little affected by the frailty distribution.

Key words: clinical trials, cross-over, Cox model, frailty, heterogeneity, omitted covariates, HIV prevention trials.

# Valorisation scientifique

## Articles

Cécilia-Joseph E., Auvert B., Broët P., Moreau T., Influence of trial duration on the bias of the estimated treatment effect in clinical trials when individual heterogeneity is ignored. *Biometrical Journal*, 2015 57 : DOI : 10.1002 / bimj.201400046.

Cécilia-Joseph E., Auvert B., Moreau T., Omitting an individual time varying binary frailty in Cox's model for analyzing clinical trials : consequences on the estimated treatment effect. En cours de soumission.

Cécilia-Joseph E., Auvert B., Moreau T., Comparison of crossover and parallel designs in clinical trials with heterogeneous population. En cours de soumission.

## Communications orales

Cécilia-Joseph E., Auvert B., Broët P., Moreau T., "Influence of trial duration on the bias of the estimated treatment effect in clinical trials when individual heterogeneity is ignored", *Journées de la Société Française de Biométrie*, Novembre 2012, Paris.

Cécilia-Joseph E., Auvert B., Moreau T., "Omitting an individual time varying binary frailty in Cox's model for analyzing clinical trials : consequences on the estimated treatment effect.", *Séminaire des doctorants LSTA Paris 6*, Février 2014, Paris.

Cécilia-Joseph E., Auvert B., Moreau T., "Comparison of crossover and parallel designs in clinical trials with heterogeneous population", *Journées du Groupe de Recherche en Statistique et Santé*, Juin 2015, Paris.

## Communication affichée

Cécilia-Joseph E., Auvert B., Moreau T., “Omitting an individual time varying binary frailty in Cox’s model for analyzing clinical trials : consequences on the estimated treatment effect.”, *Annual Conference of the International Society for Clinical Biostatistics*, Août 2013, Munich.

# Table des matières

<b>Remerciement</b>	<b>III</b>
<b>Résumé</b>	<b>V</b>
<b>Abstract</b>	<b>IX</b>
<b>Valorisation scientifique</b>	<b>XI</b>
<b>Table des matières</b>	<b>XV</b>
<b>Liste des figures</b>	<b>XX</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Contexte . . . . .	1
1.2 Le cross-over : une réponse possible au biais . . . . .	5
1.3 Objectifs et plan de la thèse . . . . .	6
<b>2 Approches existantes et caractéristiques connues du biais</b>	<b>9</b>
2.1 Notations et modèles . . . . .	10
2.1.1 Modèles spécifiques au sujet . . . . .	10
2.1.2 Modèles marginaux . . . . .	11
2.1.3 Modèle de travail . . . . .	12

2.2	Propriétés asymptotiques de l'estimateur du maximum de vraisemblance	
	partielle . . . . .	12
2.2.1	Limite asymptotique du biais . . . . .	12
2.2.2	Distribution asymptotique de l'EMV . . . . .	15
2.3	Caractéristiques connues du biais . . . . .	16
<b>3</b>	<b>Influence de la durée de l'étude sur le biais et interprétation</b>	<b>19</b>
3.1	Distributions de la fragilité . . . . .	20
3.1.1	Distribution gamma . . . . .	20
3.1.2	Distribution Log-normale . . . . .	22
3.1.3	Distribution binaire . . . . .	22
3.2	Interprétation du biais . . . . .	24
3.3	Rapport des risques marginaux . . . . .	28
3.4	Influence de la durée de l'étude sur le biais . . . . .	30
3.5	Illustration : l'essai CALGB 9633 . . . . .	37
3.6	Discussion . . . . .	42
<b>4</b>	<b>Un cas particulier de fragilité dépendante du temps</b>	<b>45</b>
4.1	Motivations . . . . .	45
4.2	Modèle de simulations . . . . .	46
4.2.1	Modèle et notations . . . . .	46
4.2.2	Génération des données . . . . .	47
4.3	Résultats . . . . .	49
4.4	Discussion . . . . .	53
<b>5</b>	<b>Utilisation d'un schéma en cross-over pour une population hétéro-</b>	
	<b>gène</b>	<b>59</b>
5.1	Adaptation des modèles au schéma en cross-over . . . . .	60

---

5.2	Comparaison avec le schéma parallèle classique . . . . .	61
5.2.1	Fragilités continues . . . . .	61
5.2.2	Fragilité catégorielle . . . . .	62
5.3	Étude du temps de changement de groupe optimal . . . . .	64
5.4	Discussion . . . . .	69
<b>Conclusion et perspectives</b>		<b>75</b>
<b>Bibliographie</b>		<b>79</b>
<b>Annexes</b>		<b>84</b>





# Table des figures

1.1	Exemple de courbes de survie avec et sans hétérogénéité, pour un nouveau traitement plus efficace que le traitement de référence. . . . .	4
1.2	Schéma en cross-over . . . . .	6
3.1	Densité de la loi gamma en fonction de la variance $\varphi$ . . . . .	21
3.2	Densité de la loi log-normale en fonction de la variance $V$ . . . . .	23
3.3	Le poids $z(t)$ en fonction de la survie marginale des 2 groupes pour les fragilités gamma, log-normale et binaire, $e^\beta = 2$ , pour un modèle exponentiel. . . . .	25
3.4	Rapport des risques marginaux en fonction de la survie marginale et de la variance de la fragilité $\varphi$ , $e^\beta = 2$ . . . . .	30
3.5	Rapport des risques marginaux en fonction de la survie marginale et de la variance de la fragilité $V$ , $e^\beta = 2$ . . . . .	31
3.6	Rapport des risques marginaux en fonction de la survie marginale et du paramètre $p$ , avec $\gamma = 1$ ou $2$ , $e^\beta = 2$ . . . . .	32
3.7	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité $\varphi$ , avec $e^\beta = 2$ . Pourcentage de censure entre parenthèse. . . . .	35

3.8	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité $V$ , avec $e^\beta = 2$ . Pourcentage de censure entre parenthèse. . . . .	35
3.9	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et $p$ , avec $e^\beta = 2$ et $\gamma = 1$ . Intervalle des pourcentages correspondant à une valeur de $\kappa$ donnée entre parenthèse. . . . .	36
3.10	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité ou du paramètre $p$ , avec $e^\beta = 1.5$ ou $2.5$ . . . . .	38
3.11	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité ou du paramètre $p$ , avec $e^\beta = 2$ . Le délai de survenue de l'évènement suivant une distribution Weibull( $k,1$ ), avec $k = 1.5, 2$ ou $2.5$ . . . . .	39
3.12	Survie théorique pour les groupes témoin et de chimiothérapie avec une durée d'étude de 7ans (a) et 11 ans (b). . . . .	41
4.1	Densité de la loi Weibull pour $CV=0.05, 1, 2$ ou $4$ et $E(V)=0.05, 0.25, 1$ ou $2$ . . . . .	50
4.2	Pourcentage du biais relatif de l'estimation de $e^\beta$ en fonction de $CV$ et $E(V)$ , pour $p = 1$ et $0.5$ , $e^\beta = 2$ , $\lambda_0 = 1$ et $T_w = 2$ . Nombre moyen de changements entre parenthèses. . . . .	51
4.3	Pourcentage du biais relatif de l'estimation de $e^\beta$ en fonction de $CV$ et $E(V)$ , pour $OR=10, 50$ et $100$ , $e^\beta = 2$ , $\lambda_0 = 1$ . . . . .	54
4.4	Pourcentage du biais relatif de l'estimation de $e^\beta$ en fonction de $CV$ et $E(V)$ , pour $p = 1$ et $0.5$ , $e^\beta = 2$ , $\lambda_0 = 1$ et $T_w = 4$ . Nombre moyen de changements entre parenthèses. . . . .	55

4.5	Pourcentage du biais relatif de l'estimation de $e^\beta$ en fonction de $CV$ et $E(V)$ , pour modèle de Weibull, paramètres de forme et d'échelle respectivement égaux à 1 et 2, $p=0.1, 0.2$ ou $0.3$ . $e^\beta = 2$ , $\lambda_0 = 1$ et $T_w = 2$ . . . . .	56
5.1	Fragilité de distribution gamma : Pourcentage du biais relatif asymptotique sur l'estimation de $\beta$ en fonction du temps de changement de groupes $T_{switch}$ , avec $T_w = 2$ . . . . .	65
5.2	Fragilité de distribution log-normale : Pourcentage du biais relatif asymptotique sur l'estimation de $\beta$ en fonction du temps de changement de groupes $T_{switch}$ , avec $T_w = 2$ . . . . .	66
5.3	Fragilité de distribution catégorielle : Pourcentage du biais relatif asymptotique sur l'estimation de $\beta$ en fonction du temps de changement de groupes $T_{switch}$ , avec $T_w = 2$ . . . . .	67
5.4	Fragilité de distribution log-normale : $T_{switch}^0$ en fonction de $T_w$ et de $\varphi$ pour $e^\beta = 2, 3$ ou $4$ . . . . .	70
5.5	Fragilité de distribution log-normale : $T_{switch}^0$ en fonction de $T_w$ et de $V$ pour $e^\beta = 2, 3$ ou $4$ . . . . .	71
5.6	Fragilité de distribution catégorielle : $T_{switch}^0$ en fonction de $T_w$ et de $P$ pour $e^\beta = 2, 3$ ou $4$ . . . . .	72
5.7	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité ou du paramètre $p$ , avec $e^\beta = 2$ . Le délai d'inclusion $T_0 = T_w/2$ ou $T_w/4$ . . . . .	84
5.8	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et de la variance de la fragilité ou du paramètre $p$ , avec $e^\beta = 2$ . Délai inclusion $T_0 = T_w/4$ dans le groupe de traitement 1 (ou 0) et $T_0 = T_w$ dans le groupe de traitement 0 (ou 1). . . . .	85

5.9	Pourcentage du biais asymptotique relatif de l'estimation de $e^\beta$ en fonction du paramètre $\kappa$ et du paramètre $p$ , avec $e^\beta = 2$ . Paramètre $\gamma = 1, 1.5, 2$ ou $3$ . . . . .	86
-----	---	----

# Chapitre 1

## Introduction

### 1.1 Contexte

Ce travail concerne la méthodologie statistique des essais thérapeutiques randomisés comparant les distributions des délais, éventuellement censurés, entre deux groupes de traitement, sous le modèle de Cox. Il porte sur les conséquences négatives de l'existence d'une hétérogénéité individuelle non prise en compte dans le modèle d'analyse de Cox, et sur la possibilité d'éviter ces conséquences.

L'essai thérapeutique randomisé a pour but d'évaluer l'efficacité d'un traitement ou intervention nouveau(elle) comparativement à un(e) autre jugé(e) de référence ou éventuellement à un placebo. Il porte sur une population de patients atteints d'une maladie donnée, ou sur un ensemble de sujets sans maladie connue. Le critère d'intérêt est, pour chaque sujet ou patient, le délai écoulé entre l'inclusion dans l'essai et l'apparition d'un événement donné qui peut être, entre autres, le décès, la guérison, la maladie ou tout autre événement lié à celle-ci. Par la suite, et pour simplifier, on parlera de décès et de temps de survie quel que soit l'évènement d'intérêt. Les groupes de traitement sont constitués par randomisation, qui consiste à répartir les sujets, au moment de leur

inclusion, de façon aléatoire équiprobable dans chacun des groupes. Cette procédure permet de réduire le biais de sélection et d'obtenir, au début du suivi, une répartition homogène des facteurs pronostiques entre les deux groupes, que ces facteurs soient connus ou non. Ainsi, on considère qu'une différence significative mise en évidence par un test de comparaison entre les deux fonctions de survie ne peut être imputable qu'au traitement.

Cependant, pour éviter un biais dans l'estimation du rapport des risques, et un manque de puissance du test de comparaison des deux distributions, la nécessité de prendre en compte les principaux facteurs pronostiques dans l'analyse statistique a largement été recommandée (Peto et al., 1977). Cette recommandation reposait sur les arguments suivants. L'homogénéité des deux groupes n'est garantie qu'au début de l'essai. Elle disparaît au cours du suivi si certains facteurs de risque sont ignorés et non pris en compte dans l'analyse. En fait, si le nouveau traitement est efficace, les sujets dont le risque est le plus élevé du fait de la valeur de leurs facteurs pronostiques présentent l'événement plus tôt dans le groupe de référence que dans celui du nouveau traitement. Par suite, la proportion de sujets à risque élevé, identique à l'origine dans les deux groupes du fait de la randomisation, diminue plus vite au cours du suivi dans le groupe de référence de sorte que le risque dans ce groupe se rapproche de celui du groupe du nouveau traitement (figure 1.1). Finalement, l'effet traitement paraît s'atténuer au cours du temps, et son estimation sera donc biaisée.

Dans le prolongement de ces constations, les conséquences de l'omission de facteurs pronostiques dans le modèle de Cox ont été l'objet de nombreux travaux, particulièrement dans le cadre des essais cliniques où plusieurs covariables sont omises (Bretagnolle et Huber-Carol, 1988), ou en précisant les caractéristiques du biais dans le cas d'une covariable oubliée de distribution particulière (Abrahamowicz et al., 2004; Chastang et al., 1988; Schmoor et Schumacher, 1997). L'existence de covariables omises dans le

modèle de Cox a pu être modélisée en introduisant un terme aléatoire individuel dans l'expression du risque instantané de chaque sujet. Ce terme de « fragilité » non observé représente les covariables oubliées et résume leur effet (Hougaard, 2000). De tels modèles avec fragilité ont été très utilisés dans un cadre différent de celui considéré ici, concernant l'analyse des données présentant un effet aléatoire partagé par plusieurs sujets au sein de sous-groupes comme des familles ou des patients provenant de mêmes structures de soin. Dans notre cadre, son utilisation pour l'analyse des données proprement dite est limitée par le fait que le terme de fragilité est individuel, non partagé par plusieurs sujets avec la conséquence que le modèle est alors non identifiable (Aalen et al., 2008). Cependant, il peut être utilisé entre autres pour calculer le biais de l'estimateur du rapport des risques, comme il sera vu par la suite. Il présente l'avantage de pouvoir considérer plusieurs covariables omises en supposant qu'elles se résument dans le terme de fragilité. Il permet aussi de traduire des types d'hétérogénéité différents de l'omission de covariables dans le modèle de Cox, comme il sera vu au Chapitre 4.

L'objectif de la thèse est de proposer l'utilisation d'un schéma de recueil des données qui permette d'éviter les inconvénients résultant de l'existence d'une hétérogénéité individuelle, comme l'obtention d'un biais dans l'estimation du rapport des risques. Ce schéma en « cross-over » est une variante de l'essai randomisé classique, où chaque sujet change de groupe de traitement après un temps fixé s'il est encore à risque à ce temps-là. Il est décrit en détail ci-après (1.2).

Ce travail a été motivé par les essais de prévention contre le VIH où l'évènement d'intérêt est la séroconversion. Dans ce type d'essais, l'existence d'une hétérogénéité inter-individuelle non observée vis-à-vis du risque d'infection est évidente. Par exemple, les facteurs de risque avérés comme le statut VIH du partenaire sexuel, le nombre de partenaires et les pratiques sexuelles, l'utilisation de préservatifs ou certains facteurs génétiques sont difficiles voire impossibles à obtenir. Les informations éventuellement



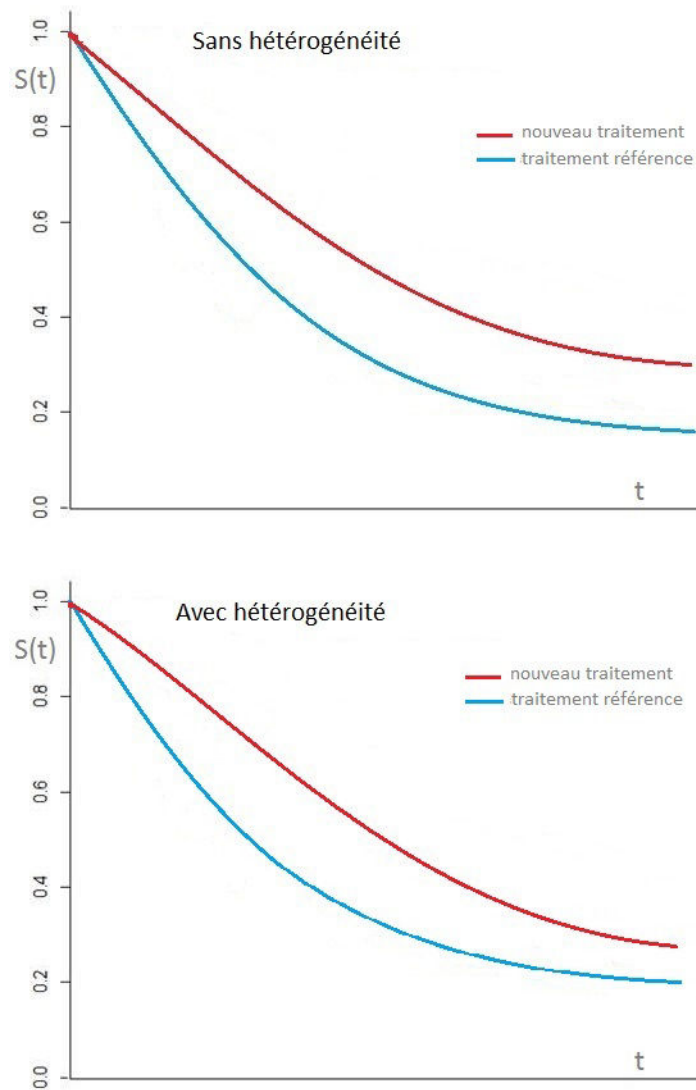


FIGURE 1.1 – Exemple de courbes de survie avec et sans hétérogénéité, pour un nouveau traitement plus efficace que le traitement de référence.

recueillies par questionnaire sur ces facteurs s'avèrent souvent peu fiables, et ils sont donc le plus souvent ignorés dans le modèle d'analyse. L'omission de ces facteurs importants peut être à l'origine d'une sous-estimation de l'effet du traitement et d'un manque de puissance du test de comparaison, ne permettant pas de valider, à tort, un nouveau traitement.

## 1.2 Le cross-over : une réponse possible au biais

Le cross-over (Gallin et al., 2003)) est un schéma d'étude dans lequel le temps de suivi des sujets est divisé en deux périodes fixées. Les sujets sont randomisés dans les deux groupes de traitement au début de l'étude et changent de groupe à la fin de la première période (figure 1.2). Dans les essais en cross-over classiques, la valeur du critère de jugement, qui peut être censurée, est mesurée à la fin de chaque période. Chaque patient a donc une mesure pour chacun des traitements. Ainsi, chaque sujet est son propre témoin et la variabilité inter-patients est supprimée et remplacée par une variabilité intra-patient, qui est souvent plus petite. Le terme de « cross-over » peut également désigner une pratique courante en oncologie consistant à attribuer le nouveau traitement, que l'on juge plus efficace, aux patients du groupe contrôle dont l'état se détériore, ce qui diffère alors du schéma croisé précédent. Dans cette thèse, on se place dans le cas d'un schéma en cross-over au sens premier du terme mais avec la particularité d'un essai pour évènement unique. Ainsi, seuls les patients n'ayant pas présenté l'évènement à la fin de la première période et n'étant pas censurés à ce temps, changent de groupe de traitement pour la seconde période. On perd alors l'avantage du sujet servant de témoin pour lui-même. Du fait de cette restriction, le schéma en cross-over est très peu utilisé dans les essais pour évènement unique. On lui préfère le schéma parallèle pour lequel les sujets suivent le même traitement tout le long du

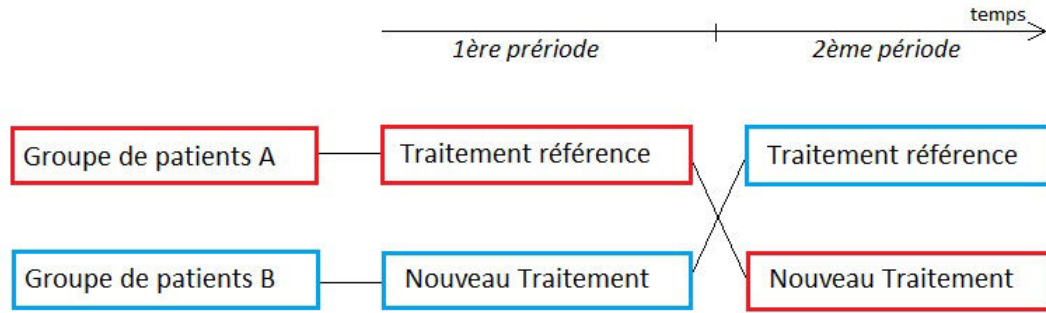


FIGURE 1.2 – Schéma en cross-over

suivi. Cependant, dans le contexte des essais de prévention à l'infection VIH, Auvert et al. (2011) ont montré que le cross-over pouvait permettre une réduction du biais sur l'estimation du rapport des risques ainsi qu'une augmentation de la puissance du test de comparaison, comparativement au schéma parallèle, en utilisant une fragilité en quatre classes. Buyze et Goetghebeur (2011) ont aussi observé les avantages du cross-over, en s'intéressant plus particulièrement à l'amélioration de l'efficacité relative du test de comparaison des deux groupes, en utilisant une fragilité gamma ou log-normale. Il s'agira donc de confirmer et préciser les résultats de ces deux travaux basées sur des études de simulations, en calculant formellement le biais asymptotique dans l'estimation de l'effet traitement.

### 1.3 Objectifs et plan de la thèse

L'objectif principal de cette thèse est de proposer l'utilisation d'un schéma de recueil des données en cross-over afin d'éviter ou d'atténuer le biais sur l'estimation du rapport des risques causé par l'omission de l'hétérogénéité individuelle dans le modèle d'analyse. L'objectif intermédiaire consiste à quantifier ce biais dans le cas de l'utilisation d'un schéma parallèle classique, tout en précisant le rôle des déterminants du biais, en

complément de résultats connus de longue date ou d'autres récemment débattus.

Le Chapitre 2 présentera les approches existantes ainsi que les caractéristiques connues du biais. Le Chapitre 3 portera sur l'influence de la durée de l'étude sur le biais en proposant une interprétation du biais. Un cas particulier de fragilité dépendante du temps, caractéristique de certains essais sur le VIH, sera étudié dans le Chapitre 4. La comparaison entre le schéma parallèle classique et le schéma en cross-over sera présentée dans le Chapitre 5. Enfin, une conclusion et quelques perspectives de travail clôtureront cette thèse.



## Chapitre 2

# Approches existantes et caractéristiques connues du biais

Dans tout ce qui suit, le modèle principalement utilisé pour exprimer l'effet de covariables oubliées est le modèle de Cox avec fragilité individuelle. Cependant, dans ce Chapitre, la formulation du modèle « complet » où figurent explicitement les covariables omises, telle qu'elle a souvent été utilisée dans la littérature, est aussi rappelée. Pour chaque approche le modèle spécifique au sujet et le modèle marginal sont précisées. Le modèle de travail mal-spécifié, commun aux deux approches, est ensuite défini. Il s'agira de calculer l'espérance et la variance asymptotiques de l'estimateur qui maximise la vraisemblance du modèle de travail, en supposant que les données suivent le modèle marginal déduit du vrai modèle spécifique au sujet. Une revue de la littérature portant sur les caractéristiques connues du biais sera également présentée, dont certains résultats s'appuient sur la formulation du biais obtenue selon ce qui précède.

## 2.1 Notations et modèles

### 2.1.1 Modèles spécifiques au sujet

Dans chacune des deux approches considérées, le groupe de traitement auquel le sujet est affecté est indiqué par une covariable observée binaire, notée  $x$ . Le délai de survenue de l'évènement est noté  $T$  et le temps de censure  $C$  est supposé indépendant de  $T$  sachant  $x$ .

Dans la première approche, on considère que les données suivent un modèle incluant, en plus de la variable indicatrice du groupe de traitement, une deuxième covariable  $z$  non-observée d'une distribution donnée. Pour chaque sujet  $i$ , la fonction de risque est alors définie par :

$$\lambda(t|z_i, x_i) = \lambda_0(t)e^{\beta x_i + \alpha z_i},$$

où  $\lambda_0(t)$  est la fonction de risque de base non spécifiée,  $\beta$  et  $\alpha$  sont les coefficients de régression inconnus correspondant respectivement à l'effet traitement et la variable d'hétérogénéité. Dans de nombreux cas et pour simplicité de calcul, la variable  $z_i$  est supposée suivre une loi de Bernoulli (Chastang et al., 1988 ; Schmoor et Schumacher, 1997), mais d'autres distributions ont parfois été considérées, par exemple gaussienne (Abrahamowicz et al., 2004).

Dans la seconde approche, qui sera celle considérée dans cette thèse, l'hétérogénéité est représentée par une variable fragilité qui est introduite multiplicativement dans le modèle de Cox. Ainsi, les données sont supposées suivre un modèle conditionnel à la fragilité tel que, pour chaque sujet  $i$ , la fonction de risque est définie par :

$$\lambda(t|w_i, x_i) = \lambda_0(t)w_i e^{\beta x_i}, \tag{2.1}$$

où  $w_i$  est la variable fragilité non observée.

Les variables  $z_i$  et  $w_i$  sont supposées indépendantes du groupe de traitement  $x_i$ .

### 2.1.2 Modèles marginaux

Pour les deux formulations, on peut définir les fonctions de survie marginales pour un sujet  $i$  par :

$$S_u(t|x_i) = \int_{-\infty}^{\infty} \{S_0(t)\}^{\exp(\beta x_i + \alpha z_i)} f(z_i) dz_i,$$

pour une covariable non-observée, et par :

$$S_u(t|x_i) = \int_0^{\infty} \{S_0(t)\}^{w_i \exp(\beta x_i)} f(w_i) dw_i,$$

pour un modèle à fragilité. Les deux expressions ne diffèrent que par un changement de variable. L'avantage de la seconde est de conduire à une expression du risque instantané à la fois générale et interprétable comme le montre le calcul ci-dessous. En écrivant :

$$S_u(t|x_i) = L_w \{ \Lambda_0(t) \exp(\beta x_i) \},$$

où  $L_w(s)$  est la transformée de Laplace de  $w$  en  $s$  et  $\Lambda_0(t)$  le risque cumulé de base de  $T$ , et compte tenu de (Aalen et al., 2008) :

$$E(w_i|T > t, x_i) = - \frac{L'_w \{ \Lambda_0(t) \exp(\beta x_i) \}}{L_w \{ \Lambda_0(t) \exp(\beta x_i) \}},$$

où  $L'_w(s)$  est la dérivée première de  $L_w(s)$ , il vient :

$$\lambda_u(t|x_i) = \lambda_0(t) e^{\beta x_i} E(w_i|T > t, x_i). \quad (2.2)$$



Il apparaît donc que le risque marginal est égal à la partie fixe du risque spécifique au sujet multiplié par l'espérance de  $w$  chez les survivants.

Dans la suite du manuscrit, les fonctions de risque et de survie marginales dans le groupe de traitement  $r$  seront respectivement notées  $\lambda_u^{(r)}(t)$  et  $S_u^{(r)}(t)$ .

### 2.1.3 Modèle de travail

Quelque soit l'approche considérée, le modèle de travail utilisé pour l'analyse des données est un modèle à risques proportionnels mal-spécifié où l'existence de la deuxième covariable pronostique, ou de la fragilité, est ignorée. Ainsi, la fonction de risque d'un sujet  $i$  est définie par :

$$\lambda_m(t|x_i) = \lambda_0(t)e^{\beta_m x_i}. \quad (2.3)$$

A noter que le coefficient  $\beta_m$  qui sera estimé avec le modèle de travail diffère du coefficient  $\beta$  intervenant dans le modèle suivi par les données. En pratique, le rapport des risques est estimé sous le modèle (2.3) par maximisation de la vraisemblance partielle.

## 2.2 Propriétés asymptotiques de l'estimateur du maximum de vraisemblance partielle

### 2.2.1 Limite asymptotique du biais

L'équation conduisant à la limite asymptotique de l'estimateur du maximum de vraisemblance partielle (EMV) du coefficient d'une covariable incluse dans le modèle de Cox quand le modèle est mal spécifié a été établie par Struthers et Kalbfleisch (1986). Elle a été généralisée à la présence de plusieurs covariables dans le modèle de travail par Bretagnolle et Huber-Carol (1988) dans la situation où le vrai modèle est un modèle

de Cox dont certaines covariables sont omises dans le modèle de travail, puis par Lin et Wei (1989) dans le cas où le vrai modèle pouvait différer d'un modèle de Cox.

Dans le cas particulier de la comparaison de deux groupes, l'équation a été précisée par Schmoor et Schumacher (1997) en se restreignant au cas où le vrai modèle serait un modèle de Cox incluant deux covariables, dont l'une serait omise dans le modèle de travail. Cette formule est identique à celle correspondant à une fragilité non observée comme considéré dans cette thèse. Cette équation, dont la solution est la limite  $\beta_m^*$  de l'EMV  $\hat{\beta}_m$  de  $\beta_m$  est donnée par :

$$\int_0^\infty \frac{\bar{y}^{(0)}(t)\bar{y}^{(1)}(t)}{\bar{y}^{(0)}(t) + \bar{y}^{(1)}(t)e^{\beta_m}} [\lambda_u^{(1)}(t) - e^{\beta_m}\lambda_u^{(0)}(t)] dt = 0 \quad (2.4)$$

où  $\bar{y}^{(r)}(t)$  est la probabilité qu'un sujet soit à risque au temps  $t$  et appartienne au groupe de traitement  $r$ . Soit  $C(t|x=r)$  la fonction de répartition du délai de censure dans le groupe de traitement  $r$ , alors :

$$\bar{y}^{(r)}(t) = P(x=r)S_u^{(r)}(t)(1 - C(t|x=r)).$$

Une justification intuitive de l'équation (2.4) est la suivante. Le score calculé sous le modèle de travail est donné par :

$$U(\beta_m) = \sum_{i=1}^n \int_0^\infty \left( x_i - \frac{n_1(t)e^{\beta_m}}{n_0(t) + n_1(t)e^{\beta_m}} \right) dN_i(t),$$

qui peut aussi s'écrire :

$$U(\beta_m) = \sum_{i=1}^n \int_0^\infty \frac{n_0(t)n_1(t)}{n_0(t) + n_1(t)e^{\beta_m}} \left( \frac{x_i}{n_1(t)} - \frac{(1-x_i)e^{\beta_m}}{n_0(t)} \right) dN_i(t), \quad (2.5)$$

où  $n_r(t)$  est le nombre de sujets à risque dans le groupe de traitement  $r$  au temps  $t$ , et  $N_i(t) := \mathbf{1}_{\{T_i \leq t, \delta_i=1\}}$ . Il apparaît alors que le membre de gauche de (2.4) n'est autre

que la limite en probabilité du score mis sous la forme ci-dessus lorsque le vrai modèle est le modèle marginal défini par  $\lambda_u^{(r)}(t)$ ,  $r = 0, 1$ .

A partir de la même observation, Struthers et Kalbfleisch (1986) avaient conclu de façon équivalente que le membre de gauche de l'équation (2.4) est obtenu par dérivation de la limite en probabilité de la log-vraisemblance partielle calculée sous le modèle de travail à risques proportionnels, alors que le vrai modèle est le modèle marginal déduit du modèle complet spécifique au sujet. Dans l'approche utilisée par ces auteurs, le modèle complet était le modèle de Cox incluant une covariable non-observée.

On peut également remarquer que, si le modèle de travail est le vrai modèle, alors les risques instantanés marginaux  $\lambda_u^{(r)}(t)$  dans (2.4) sont identiques aux risques instantanés définissant le modèle de travail et l'équation (2.4) est vérifiée lorsque  $\beta_m$  prend la valeur de l'EMV  $\hat{\beta}_m$  comme attendu.

Pour exemple, avec l'approche reposant sur l'oubli de covariables dans le modèle de Cox, et en supposant que le vrai modèle comprend une seule covariable  $x$  de Bernoulli en plus de la variable traitement, les éléments nécessaires au calcul de (2.4) sont donnés par (Schmoor et Schumacher ; 1997) :

$$\bar{y}^{(r)}(t) = P(x = r) \{ (1 - p) S(t|x_i = r, z_i = 0) + p S(t|x_i = r, z_i = 1) \} \{ 1 - C(t) \}$$

et

$$\lambda_u^{(r)}(t) = \frac{(1 - p) \lambda(t|x_i = r, z_i = 0) S(t|x_i = r, z_i = 0) + p \lambda(t|x_i = r, z_i = 1) S(t|x_i = r, z_i = 1)}{(1 - p) S(t|x_i = r, z_i = 0) + p S(t|x_i = r, z_i = 1)}.$$

### 2.2.2 Distribution asymptotique de l'EMV

La distribution de  $\sqrt{n}(\hat{\beta}_m - \beta_m^*)$  est asymptotiquement normale centrée en zéro et dont la variance peut être estimée par un estimateur de type « sandwich » sous le modèle 2.2 convergeant en probabilité vers  $V^* = W(\beta_m^*)/i(\beta_m^*)^2$  (Lin et Wei, 1989 ; Schmoor et Schumacher, 1997), où  $W(\beta_m^*)$  est la variance asymptotique de la statistique du score et  $i(\beta_m^*)$  est la limite stochastique de l'opposé de la dérivée seconde de la log-vraisemblance divisé par  $n$ . Elles sont définies par les expressions suivantes :

$$W(\beta_m^*) = \int_0^\infty \frac{\bar{y}^{(0)}(t)\bar{y}^{(1)}(t)}{[\bar{y}^{(0)}(t) + \bar{y}^{(1)}(t)e^{\beta_m^*}]^2} [\bar{y}^{(0)}(t)\lambda_u^{(1)}(t) + e^{2\beta_m^*}\bar{y}^{(1)}(t)\lambda_u^{(0)}(t)] dt$$

$$i(\beta_m^*) = \int_0^\infty \frac{\bar{y}^{(0)}(t)\bar{y}^{(1)}(t)e^{\beta_m^*}}{[\bar{y}^{(0)}(t) + \bar{y}^{(1)}(t)e^{\beta_m^*}]^2} [\bar{y}^{(0)}(t)\lambda_u^{(0)}(t) + \bar{y}^{(1)}(t)\lambda_u^{(1)}(t)] dt.$$

Dans le cas où les données suivraient effectivement un modèle à risques proportionnels,  $W(\beta_m^*) = i(\beta_m^*)$  et la variance asymptotique est égale à  $i(\beta_m^*)^{-1}$ . On obtient alors l'expression habituellement utilisée pour calculer l'estimation de la variance de l'EMV, c'est-à-dire l'inverse de l'information de Fisher.

Dans le cas où le modèle de travail est mal spécifié, la variance asymptotique est calculée grâce à l'expression « sandwich »  $V^*$  qui est une estimation robuste de la variance et ce même lorsque le modèle de travail est incorrect.

Les propriétés asymptotiques de  $\hat{\beta}_m$  présentées ici sont valables quelque soit l'approche utilisée, covariable ou fragilité omise. Elles ont été utilisées dans la littérature en considérant une covariable oubliée dans le modèle de Cox afin de définir plusieurs caractéristiques du biais qui sont présentées dans la section suivante.

## 2.3 Caractéristiques connues du biais

Dans un premier temps, Bretagnolle et Huber-Carol (1988) se sont placés dans un cadre très général en montrant que l’omission de covariables dans le modèle de Cox entraînait une sous-estimation en valeur absolue du coefficient de régression des covariables incluses, qui existe même si les covariables omises sont indépendantes des covariables incluses dans le modèle. Puis, de nombreux travaux viennent préciser ce résultat en étudiant les caractéristiques du biais dans le cadre d’essais randomisés comparant deux groupes de traitement via différentes approches.

Dans ce contexte et dans le cas d’une distribution binaire de l’hétérogénéité, plusieurs facteurs du biais ont été identifiés. Dans un cadre paramétrique, Gail et al. (1984) ont calculé le biais asymptotique en utilisant la méthode des moments, tandis que Chastang et al. (1988) ont effectué des études de simulation. Une association positive entre le biais et la vraie valeur du risque relatif, ainsi qu’avec le coefficient lié à la covariable oubliée est alors observée. Ce résultat est retrouvé dans le cadre semi-paramétrique notamment par Schumacher et al. (1987) et Schmoor et Schumacher (1997) s’intéressant au biais asymptotique en s’appuyant sur l’équation (2.4) présentée à la section 2.2. Tous ces travaux constatent également un impact de la proportion de sujets « fragiles » sur l’estimation du risque relatif, avec un biais maximal lorsque cette proportion est proche de 0.5. Dans le cadre semi-paramétrique, qui correspond au cadre de ce travail, une diminution du biais est observée lorsque la proportion de censure augmente. Notons tout de même que cette relation n’est pas forcément vérifiée dans le cadre paramétrique comme l’observent par exemple Chastang et al. avec un modèle multiplicatif exponentiel.

Une autre conséquence de l’omission de l’hétérogénéité dans le modèle d’analyse est une diminution de la puissance du test de comparaison. Ce problème est aussi évoqué dans les travaux cités ci-dessus. A partir de formules du biais et de la variance

asymptotiques, Schmoor & Schumacher (1997) soulignent le fait que le manque de puissance est principalement causé par le biais, la variance n'étant que très faiblement surestimée. Aussi, Lagakos & Schoenfeld (1984) se sont particulièrement intéressés à la puissance en mettant en avant une baisse de l'efficacité asymptotique du test du score lorsqu'on s'éloigne de la proportionnalité. Betensky et al. (2002) ont confirmé ces résultats théoriques en s'intéressant à la puissance et à la taille d'échantillon pour un essai réel en oncologie, considérant l'omission d'une hétérogénéité moléculaire des patients.

D'autre part, Lin et al. (2013) ont proposé une autre approche basée sur une analyse de sensibilité permettant d'évaluer le biais causé par l'omission de covariables dans le modèle de Cox, applicable dans les essais randomisés ou dans les études observationnelles.



## Chapitre 3

# Influence de la durée de l'étude sur le biais et interprétation

L'amplitude du biais du rapport des risques en présence d'hétérogénéité non prise en compte dépend de différents facteurs qui ont été identifiés et étudiés dans de nombreux articles, comme il a été vu plus haut (§2.3). Aucun résultat ne porte sur la durée de l'essai, bien que son implication soit suggérée par le fait suivant (Xu & O'Quigley ; 2000). Dans la situation où le coefficient d'une covariable incluse dans un modèle de Cox serait en réalité variable au cours du temps, et non constant comme le suppose le modèle utilisé, l'estimation de ce coefficient supposé constant converge vers une moyenne des « vrais » coefficients au cours du temps. Dans le cadre étudié ici, ce résultat devrait avoir pour conséquence que le rapport des risques estimé à partir du modèle de travail convergerait vers une moyenne au cours du temps des rapports des risques du modèle marginal en chaque temps. Par suite, le biais sur l'estimation du rapport des risques issue du modèle de travail augmenterait avec la durée de l'essai. Cette hypothèse est envisagée dans la suite de ce chapitre. L'interprétation du biais comme ci-dessus est d'abord précisée, puis la relation entre le biais et la durée de l'essai est évaluée par



calcul pour différentes situations.

## 3.1 Distributions de la fragilité

Dans ce chapitre, trois distributions de fragilité sont utilisées, gamma, log-normale et binaire.

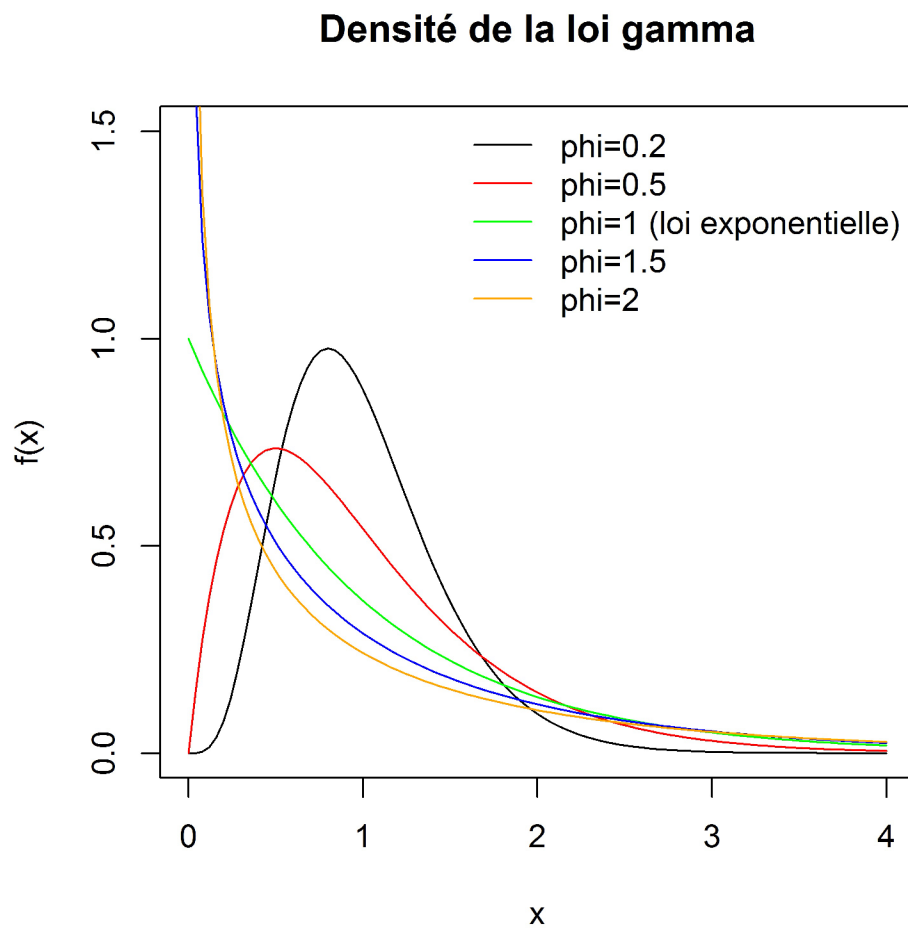
### 3.1.1 Distribution gamma

La fonction de densité de la loi gamma est définie sur  $(0, \infty)$  par :

$$f(x, k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\Gamma(k) \theta^k},$$

avec les paramètres de forme  $k$  et d'échelle  $\theta$  strictement positifs. Dans la suite, ils sont définis de façon à ce que l'espérance et la variance de la variable fragilité soient respectivement égales à 1 et à une valeur  $\varphi$  fixée, c'est-à-dire  $k = 1/\varphi$  et  $\theta = \varphi$ . Pour  $\varphi = 1$ , on retombe sur la densité de la loi exponentielle. La figure 3.1 représente la fonction de densité de la loi gamma pour différentes valeurs de la variance  $\varphi$ .

Sans perte de généralité, l'espérance de la fragilité peut être supposée égale à 1 (Aalen et al., 2008). Pour une valeur de  $w_i$  inférieure ou supérieure à 1, le sujet  $i$  sera moins fragile, respectivement plus fragile, que la moyenne. L'hétérogénéité de la population est alors quantifiée par la variance  $\varphi$ . La distribution Gamma est souvent choisie dans la littérature en raison de sa maniabilité. Aussi, elle permet d'avoir des expressions explicites simples des fonctions d'intérêt. En plus de ses propriétés mathématiques, Abbring et Van Der Berg (2007) apportent des arguments théoriques quant à l'utilisation d'une distribution gamma pour la fragilité. En effet, ils montrent la convergence de la distribution de l'hétérogénéité des survivants vers une loi gamma, dans le cas où cette distribution serait définie sur  $(0, \infty)$  et aurait une « variation régulière » en 0 avec un

FIGURE 3.1 – Densité de la loi gamma en fonction de la variance  $\varphi$ .

exposant  $-\infty < \rho < \infty$ , selon la définition de Feller (1971, §VIII-8). Parmi ce type de distribution, on peut noter les fonctions exponentielle, uniforme ou encore normale tronquée. Cependant, à distance finie, il a été montré que l'estimation des coefficients des covariables incluses dans le modèle, dépendait fortement de la distribution de la fragilité (Keiding et al., 1997).

### 3.1.2 Distribution Log-normale

Une variable  $X$  suit une distribution log-normale de paramètres  $\mu$  et  $\sigma^2$  si la variable  $\log(X)$  suit une distribution normale d'espérance  $\mu$  et de variance  $\sigma^2$ . La fonction de densité de la loi log-normale est définie sur  $(0, \infty)$  par :

$$f(x, \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right).$$

Comme pour la fragilité gamma, les paramètres sont définis tels que l'espérance et la variance de la fragilité soient respectivement égales à 1 et une valeur  $V$  fixée. Ainsi,  $\mu = -\log(V + 1)/2$  et  $\sigma^2 = \log(V + 1)$ . La figure 3.2 représente la fonction de densité de la loi log-normale pour différentes valeurs de la variance  $V$ . Si on suppose que plusieurs covariables ont pu être oubliées dans l'analyse, il est usuel de penser que la combinaison linéaire de celle-ci telle qu'elle existe dans la « vraie » fonction de risque peut être représentée par une distribution gaussienne.

### 3.1.3 Distribution binaire

On suppose que la variable fragilité  $w_i$  peut prendre les valeurs 1 ou  $e^\gamma$  avec les probabilités  $p$  ou  $1 - p$ , respectivement. Dans le cas où  $\gamma = 1$ , la variable  $\log(W)$  suit alors une loi de Bernoulli de paramètre  $p$ . La distribution binaire est naturellement utilisée dans le cas où on considère une unique covariable oubliée dans le modèle, la

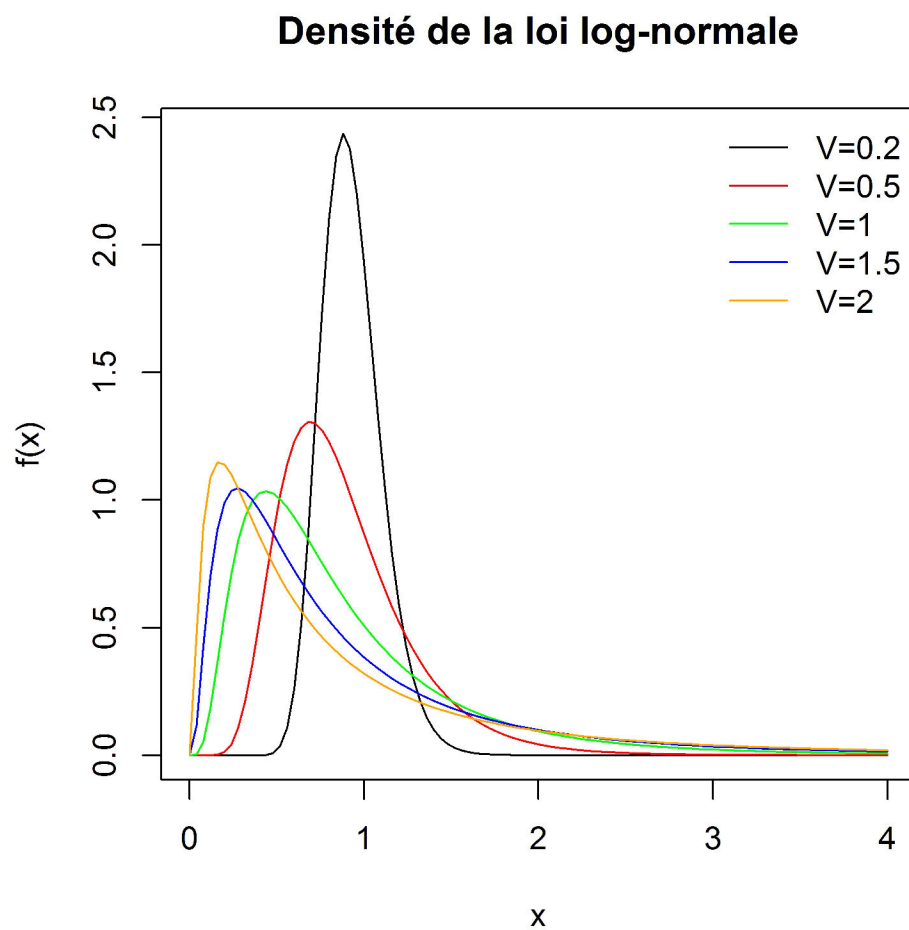


FIGURE 3.2 – Densité de la loi log-normale en fonction de la variance  $V$ .

valeur du paramètre  $\gamma$  permettant de quantifier l'influence de cette covariable sur la survie.

Dans la suite du manuscrit, si aucune précision n'est donnée, la distribution binaire considérée comprendra une valeur  $\gamma = 1$ .

## 3.2 Interprétation du biais

La propriété déclinée ici concernant la limite de l'estimation du rapport des risques est similaire à celle montrée par Xu et O'Quigley (2000) pour un coefficient de régression d'un modèle de Cox mal-spécifié où le vrai coefficient varie au cours du temps. Plus précisément, on peut montrer que la limite  $\exp(\beta_m^*)$  obtenue à partir de l'équation (2.4) est une moyenne pondérée des rapports des risques marginaux au cours du temps. En supposant que le rapport des risques  $\lambda_u^{(1)}(t)/\lambda_u^{(0)}(t)$  soit égal à  $\exp(\beta_m)$  à chaque temps  $t$ , ce qui serait requis si le modèle (2.3) était le vrai modèle, sa valeur serait aussi donnée par  $\int_0^\infty \lambda_u^{(1)}(t)dt / \int_0^\infty \lambda_u^{(0)}(t)dt = \exp(\beta_m)$ . En pondérant par le poids  $z(t)$ , on obtient :

$$\int_0^\infty z(t)\lambda_u^{(1)}(t)dt = \int_0^\infty e^{\beta_m} z(t)\lambda_u^{(0)}(t)dt,$$

qui correspond à l'équation (2.4) lorsqu'on choisit  $z(t) = \bar{y}^{(0)}(t)\bar{y}^{(1)}(t) / (\bar{y}^{(0)}(t) + \bar{y}^{(1)}(t)e^{\beta_m})$ .

Les poids  $z(t)$ , qui dépendent de  $\beta_m$ , ont été calculés dans différentes situations, avec le délai de survenue de l'évènement qui suit soit une distribution exponentielle (figure 3.3) ou Weibull, et la fragilité suivant une distribution Gamma, Log-Normale ou binaire. Il apparaît que le poids  $z(t)$  diminue régulièrement vers zéro, en fonction du temps, dans tous les cas considérés, ce qui indique que les rapports des risques précoces ont plus d'impact sur le calcul de  $\beta_m^*$  que les rapports des risques tardifs.

Une même interprétation serait obtenue en utilisant une autre approche partant de la dérivée de l'espérance du score  $U(\beta_m)$  dans (2.5). Tout d'abord, on peut noter que

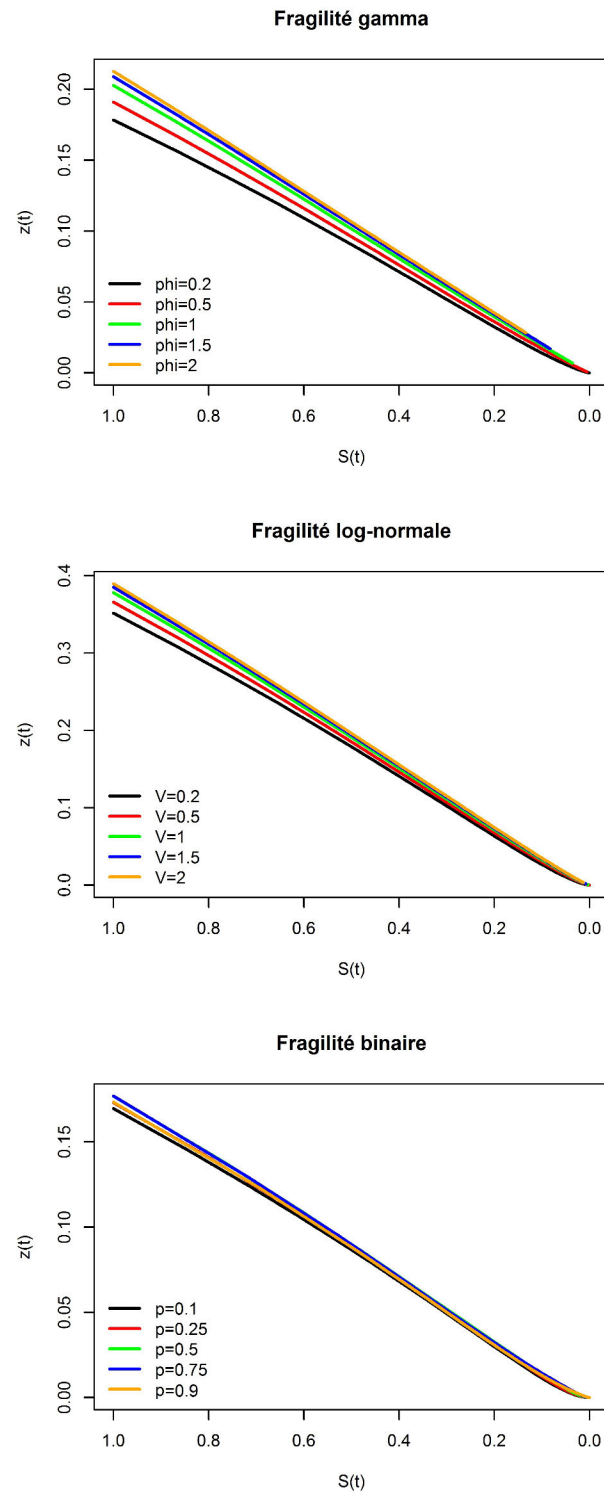


FIGURE 3.3 – Le poids  $z(t)$  en fonction de la survie marginale des 2 groupes pour les fragilités gamma, log-normale et binaire,  $e^\beta = 2$ , pour un modèle exponentiel.

pour un sujet  $i$  dans le groupe de traitement 1 (respectivement 0) et en notant  $Y_i(t)$  la variable indicatrice du statut « à risque » du sujet  $i$  au temps  $t$ , on a :

$$x_i E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 1\} = x_i \lambda_u^{(1)}(t) Y_i(t) dt,$$

et

$$(1 - x_i) E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 0\} = (1 - x_i) \lambda_u^{(0)}(t) Y_i(t) dt,$$

où  $\mathcal{F}_{t-}$  est la sigma-algèbre du passé jusqu'à  $t^-$ , et l'espérance sachant ( $\mathcal{F}_{t-}$ ) est sous le modèle (2.3). Il s'en suit qu'en sommant sur  $i$  l'expression ci-dessus et en utilisant  $\lambda_u^{(r)}(t)$  pour désigner la valeur commune de  $\lambda_u(t | x_i = r)$  pour tous les sujets  $i$  qui appartiennent au groupe  $r$ , le rapport des risques marginaux au temps  $t$  entre les groupes 1 et 0 peut s'écrire :

$$\frac{\lambda_u^{(1)}(t)}{\lambda_u^{(0)}(t)} = \frac{\sum_i x_i E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 1\} / n_1(t)}{\sum_i (1 - x_i) E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 0\} / n_0(t)}.$$

En supposant que ce rapport soit constant au cours du temps et égal à  $\exp(\beta_m)$  comme stipulé par le modèle (2.3), et en intégrant sur  $t$  après avoir pondéré par les poids  $w(t)$ , il s'en suit que :

$$\begin{aligned} \int_0^\infty w(t) \sum_i x_i E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 1\} / n_1(t) \\ - e^{\beta_m} \int_0^\infty w(t) \sum_i (1 - x_i) E \{dN_i(t) | \mathcal{F}_{t-}; x_i = 0\} / n_0(t) = 0. \end{aligned}$$

Enfin, en choisissant  $w(t) = n_0(t)n_1(t) / [n_0(t) + n_1(t)\exp(\beta_m)]$ , le membre de gauche de l'expression ci-dessus est l'espérance du score  $U(\beta_m)$  sachant ( $\mathcal{F}_{t-}$ ), qui sous le modèle (2.3), converge en probabilité vers le membre de gauche de l'équation (2.4).

Xu et O'Quigley (2000) ont obtenu un résultat similaire qui porte sur le logarithme  $\beta_m$  du rapport des risques mais nécessite que le modèle dépendant du temps soit proche du modèle à risques proportionnels. En adaptant les calculs des auteurs à notre situation et en posant :

$$\beta(t) = \log \left\{ \frac{\lambda_u^{(1)}(t)}{\lambda_u^{(0)}(t)} \right\},$$

une approximation de Taylor du premier ordre de l'intégrale de (2.4) autour de  $\beta_m = \beta(t)$  conduit à :

$$\int_0^\infty v(t) \left\{ \bar{y}^{(1)}(t) \lambda_u^{(1)}(t) + \bar{y}^{(0)}(t) \lambda_u^{(0)}(t) \right\} (\beta(t) - \beta_m) dt = 0$$

avec

$$v(t) = \frac{\bar{y}^{(0)}(t) \bar{y}^{(1)}(t) e^{\beta(t)}}{\left\{ \bar{y}^{(0)}(t) + \bar{y}^{(1)}(t) e^{\beta(t)} \right\}^2}.$$

Il s'en suit que  $\beta_m^*$  est approximativement égal à :

$$\frac{\int_0^\infty v(t) \left\{ \bar{y}^{(1)}(t) \lambda_u^{(1)}(t) + \bar{y}^{(0)}(t) \lambda_u^{(0)}(t) \right\} \beta(t) dt}{\int_0^\infty v(t) \left\{ \bar{y}^{(1)}(t) \lambda_u^{(1)}(t) + \bar{y}^{(0)}(t) \lambda_u^{(0)}(t) \right\} dt},$$

qui apparait comme une moyenne pondérée de  $\beta(t)$  avec les poids  $v(t)$ . Dans la suite de ce chapitre, le rapport des risques est préféré à son logarithme afin d'éviter l'approximation mentionnée ci-dessus.

En conséquence, d'après cette interprétation, les variations du rapport des risques marginaux en fonction du temps donnent une indication sur les variations de  $\exp(\beta_m)$  en fonction de la durée de l'étude. Elles font l'objet du paragraphe suivant.



### 3.3 Rapport des risques marginaux

D'après le modèle (2.2) le rapport des risques marginaux est égal à :

$$RR(t) = e^{\beta} \frac{E(w_i|T > t, x_i = 1)}{E(w_i|T > t, x_i = 0)}.$$

Pour une distribution Gamma d'espérance 1 et de variance  $\varphi$ , l'espérance de la fragilité sachant  $T > t$  est donnée par :

$$E(w_i|T > t, x_i) = \frac{1}{1 - \varphi e^{\beta x_i} \log(S_0(t))}.$$

Dans le cas d'une variable de fragilité binaire, elle s'écrit :

$$E(w_i|T > t, x_i) = \frac{pe^{\gamma} S_0(t)^{\exp(\gamma + \beta x_i)} + (1 - p) S_0(t)^{\exp(\beta x_i)}}{p S_0(t)^{\exp(\gamma + \beta x_i)} + (1 - p) S_0(t)^{\exp(\beta x_i)}}.$$

Pour la fragilité log-normale, le rapport des risques marginaux ne peut pas être exprimé sous une forme analytique et doit être calculé numériquement.

Les figures 3.4, 3.5 et 3.6 représentent les variations du rapport des risques marginaux en fonction de la survie marginale pour les fragilités gamma, log-normale et binaire pour  $\gamma = 1$  ou 2, avec un effet traitement  $e^{\beta} = 2$ .

Dans le cas des fragilités continues, le rapport des risques diminue régulièrement au cours du temps à partir de la vraie valeur de l'effet traitement. Cette diminution est d'autant plus rapide que la variance de la fragilité est élevée. Aussi, pour la fragilité gamma, on peut voir que le rapport des risques tend vers 1 lorsque la survie tend vers 0. Autrement dit, l'effet du traitement chez les survivants s'estompe jusqu'à devenir pratiquement nul. Pour la fragilité log-normale, cette diminution est moindre mais le rapport des risques atteint tout de même la valeur 1.3.

Dans le cas de la fragilité binaire, le rapport des risques diminue régulièrement au

cours du temps jusqu'à atteindre un minimum, puis réaugmente lorsque la survie tend vers 0, jusqu'à sa valeur initiale, c'est-à-dire la vraie valeur de l'effet traitement. Plus la probabilité  $p$  est élevée, plus ce minimum est atteint précocement. Ce dernier est plus élevé pour  $\gamma = 1$  par rapport à  $\gamma = 2$  et ainsi que pour les plus petites valeurs de  $p$ .

Ces résultats sont en adéquation avec les observations faites par Schumacher et al. (1987) sur l'évolution du rapport des risques au cours du temps dans le cas binaire sous l'approche d'une covariable oubliée, et par Aalen et al. (2008) dans le cas d'une fragilité gamma.

La randomisation au début du suivi est censée assurer une répartition homogène des sujets dans les deux groupes de traitement. Les plus « fragiles », c'est-à-dire ceux ayant une fragilité égale à  $e^\gamma$  dans le cas d'une fragilité binaire, ont tendance à présenter l'évènement plus précocement dans le groupe de traitement 1 comparativement au groupe de traitement 0. Cela crée donc une différenciation de la composition des deux groupes au cours du temps, entraînant une diminution du rapport des risques. Cependant, dans le cas binaire, quand la survie tend vers 0, la proportion des sujets les moins « fragiles », donc ayant une fragilité de 1, tend vers 1, et ce dans les deux groupes de traitement. Par conséquent, on tend vers une homogénéisation des deux groupes, et le rapport des risques tend alors de nouveau vers la vraie valeur de l'effet traitement. Dans les cas continus, l'infinité de valeurs pouvant être prises par la fragilité ne permet pas d'observer ce phénomène.

D'après les variations du rapport des risques marginaux en fonction du temps, on peut anticiper une diminution de l'estimation de  $\exp(\beta_m)$  avec la durée de l'étude.

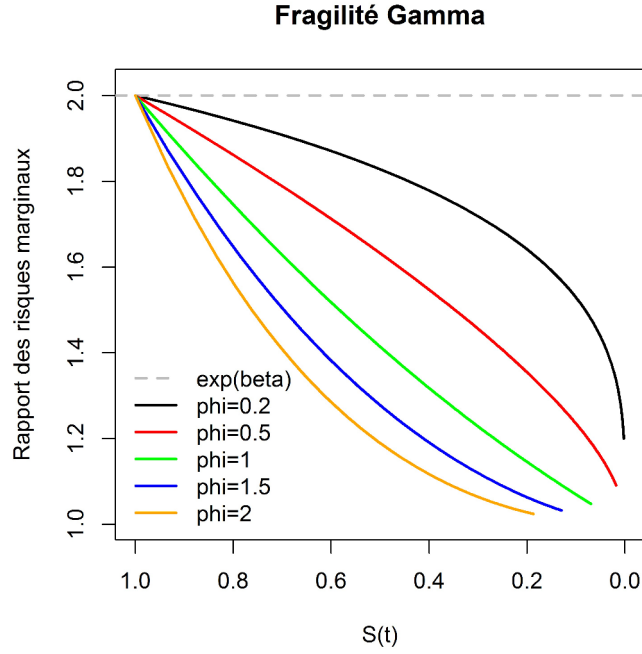


FIGURE 3.4 – Rapport des risques marginaux en fonction de la survie marginale et de la variance de la fragilité  $\varphi$ ,  $e^\beta = 2$ .

### 3.4 Influence de la durée de l'étude sur le biais

Dans cette section, l'évolution du biais relatif asymptotique (« BR »), défini par  $(\exp(\beta_m^*) - \exp(\beta)) / \exp(\beta)$ , est exprimé en fonction de la durée de l'étude. Plusieurs hypothèses sont posées concernant les distributions de la fragilité, du délai de survenue de l'évènement et du schéma de censure.

La variance des fragilités continues prend les valeurs 0.2, 0.5, 1, 1.5 ou 2. Le paramètre  $\gamma$  de la fragilité binaire prend les valeurs 1, 1.5, 2 ou 3, et le paramètre  $p$  vaut 0.1, 0.25, 0.5, 0.75 ou 0.9. Le délai de survenue de l'évènement suit une distribution exponentielle avec un risque de base égal à 1 et un effet traitement égal à 1.5, 2 ou 2.5. Une distribution Weibull est également étudiée avec un paramètre d'échelle fixé à 1 et un paramètre de forme  $k$  prenant les valeurs 1.5, 2 ou 2.5.

Il faut noter que l'intégrale dans l'équation (2.4) dépend de la distribution de cen-

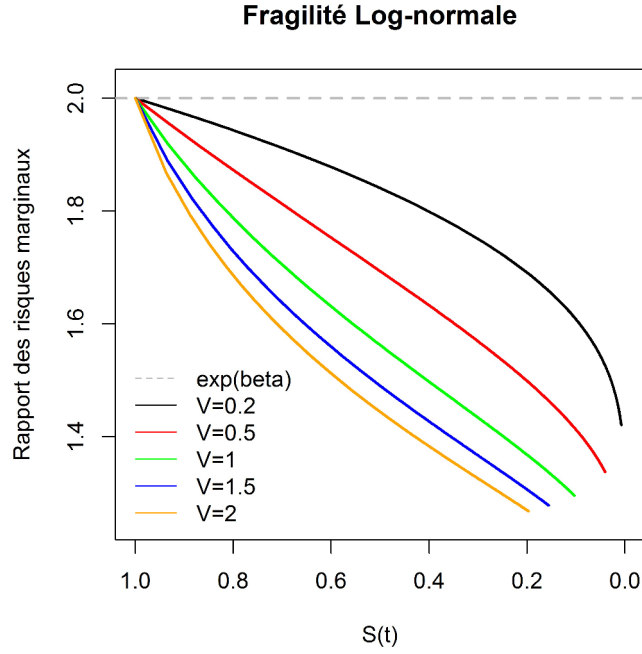


FIGURE 3.5 – Rapport des risques marginaux en fonction de la survie marginale et de la variance de la fragilité  $V$ ,  $e^\beta = 2$ .

sure. Dans un premier temps, les sujets sont supposés être inclus dans l'étude de façon uniforme, à partir d'une origine de temps jusqu'à un temps donné  $T_0$ , alors que la durée totale de l'étude  $T_w$  est la même pour tous les sujets, comme il est souvent considéré dans les essais thérapeutiques randomisés. Par conséquent, le temps de censure est indépendant du groupe de traitement et est uniformément distribué sur l'intervalle de temps  $(T_w - T_0, T_w)$ . A noter que  $C(t) = 1$  pour  $t > T_w$  et (2.4) n'est pas définie au delà de  $T_w$ . Dans ce cadre, l'influence de la durée de l'étude sur le biais peut être exprimée en reformulant l'équation (2.4) :

$$\int_0^{T_w} \frac{(T_w - t)}{T_0} \frac{P(x=0)S_u^{(0)}(t)P(x=1)S_u^{(1)}(t)}{P(x=0)S_u^{(0)}(t) + P(x=1)S_u^{(1)}(t)e^{\beta_m}} [\lambda_u^{(1)}(t) - e^{\beta_m}\lambda_u^{(0)}(t)] dt = 0. \quad (3.1)$$

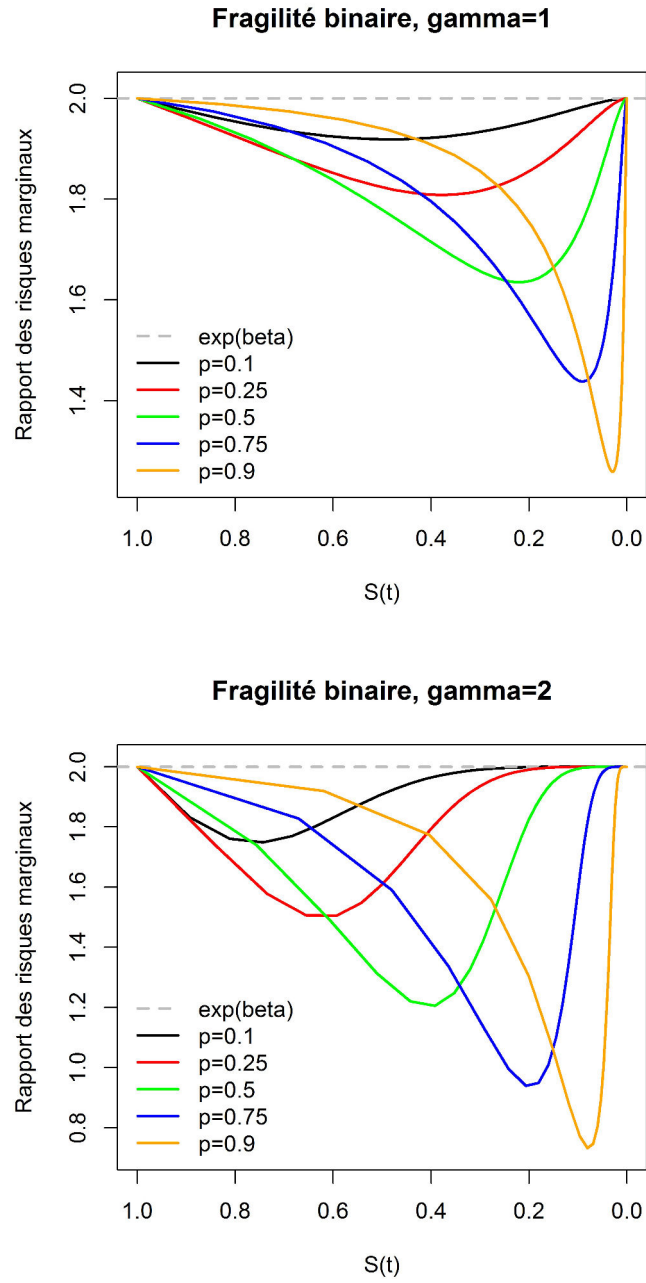


FIGURE 3.6 – Rapport des risques marginaux en fonction de la survie marginale et du paramètre  $p$ , avec  $\gamma = 1$  ou  $2$ ,  $e^\beta = 2$ .

Trois cas sont considérés :  $T_0 = T_w$ ,  $T_0 = T_w/2$ , et  $T_0 = T_w/4$ .

Il peut arriver que dans certains cas,  $T_w$  diffère d'un groupe à l'autre. Ainsi, la distribution de censure dépend du groupe de traitement. Cette situation est considérée dans un second temps.

Les figures 3.7, 3.8 et 3.9 montrent l'évolution du BR en fonction de la durée de l'étude dans le cas où le délai de survenue de l'évènement suit une distribution exponentielle, avec un effet traitement de 2 et un risque de base de 1, pour les fragilités gamma, log-normale et binaire avec  $\gamma = 1$ . La durée de l'étude est exprimée à l'aide du paramètre  $\kappa$  tel que :

$$\kappa = \frac{T_w}{\text{Médiane de survie}},$$

où la médiane de survie est calculée à partir du mélange des distributions du délai d'évènement dans les deux groupes de traitement, et  $\kappa$  vaut 0.5, 1, 2, 4, 6 ou 8. Le biais obtenu pour une durée d'étude  $T_w$  donnée est fortement dépendant de la valeur de la survie de l'échantillon à  $T_w$ , comme le laisse supposer les variations des rapports des risques marginaux à la section 3.3, et donc a fortiori des valeurs choisies pour les différents paramètres jouant sur cette survie. Utiliser la médiane de survie comme échelle pour la durée de l'étude permet alors d'avoir des résultats généraux et interprétables. C'est pourquoi le rapport  $\kappa$  a été préféré comme unité de temps plutôt que la durée d'étude propre.

Pour les fragilités continues, les variations du BR sont présentées selon la variance de la fragilité, et selon le paramètre  $p$  pour la fragilité binaire. Pour chaque cas de figure,  $T_0 = T_w$ , et le pourcentage de censure est calculé selon la distribution de fragilité, la valeur de  $\varphi$  ou  $V$ , ou  $p$ , et la valeur  $\kappa$ .

Les figures 3.7, 3.8 et 3.9 montrent que pour chaque distribution de fragilité, le BR augmente en valeur absolue avec la durée de l'étude et est sensiblement plus élevé pour

les fragilités continues par rapport à la fragilité binaire. Pour une durée d'étude donnée et pour les fragilités continues, il augmente également avec la valeur de la variance de la fragilité, alors que pour la fragilité binaire il augmente avec la valeur de  $p$  jusqu'à  $p = 0.5$  et diminue pour les plus grandes valeurs de  $p$ . En effet, pour une étude qui s'arrête à  $T_w = 2 * \text{médiane de survie}$ , le BR varie de 4.7% pour  $\varphi = 0.2$  à 25.0% pour  $\varphi = 2$  dans le cas d'une fragilité gamma, de 4.2% pour  $V = 0.2$  à 15.9% pour  $V = 2$  dans le cas d'une fragilité log-normale, et n'excède pas 5.8% qu'importe la valeur de  $p$  pour une fragilité binaire.

Notons que le biais dépend aussi du pourcentage de censure qui diminue régulièrement lorsque  $\kappa$  augmente et, dans une moindre mesure, lorsque la variance de la fragilité diminue, comme le montrent les figures 3.7 et 3.8. Dans le cas d'une fragilité binaire et pour un  $\kappa$  donné, ce pourcentage augmente quand la valeur de  $p$  diminue de 1 à 0.5 puis augmente avec des variations qui restent modérées. C'est pourquoi seules les limites des intervalles de ces pourcentages pour chaque valeur de  $\kappa$  ont été représentées sur la figure 3.9. Par conséquent, pour une variance de la fragilité donnée ou pour une valeur de  $p$  donnée, le biais augmente en valeur absolue lorsque le pourcentage de censure déduit de  $\kappa$  diminue. Pour une durée d'étude donnée, il augmente avec le pourcentage de censure associé à la variance de la fragilité ou à la valeur  $p$ .

D'autres résultats sont obtenus pour différentes valeurs du rapport des risques (figure 3.10) et lorsque l'on considère une distribution de Weibull pour le délai de survenue de l'évènement (figure 3.11). Comme on peut le voir sur les figures 3.10 et 3.11, les mêmes tendances sont observées. Avec une distribution de Weibull, on obtient des valeurs similaires du biais en fonction du paramètre  $\kappa$  et de la variance de la fragilité ou la valeur de  $p$ . Sur la figure 3.10, on peut noter une association positive entre le biais et la vraie valeur de l'effet traitement  $e^\beta$ . Les variations du BR dans le cas où le temps d'inclusion des sujets  $T_0$  est égal à  $T_w/2$  ou  $T_w/4$  sont présentées dans l'Annexe A. Les

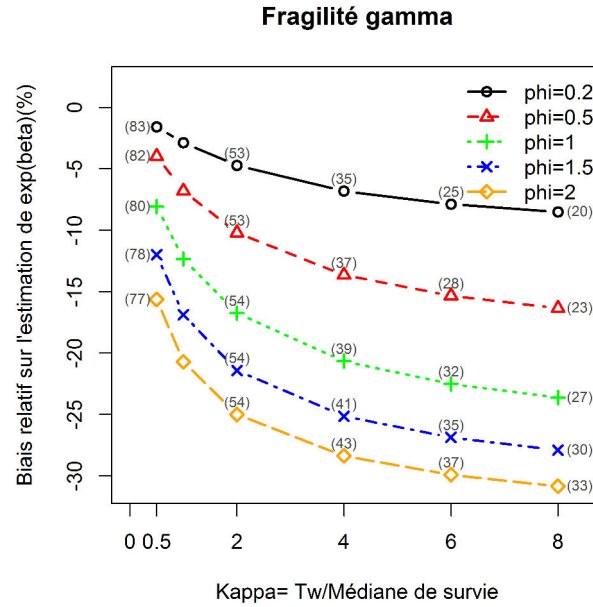


FIGURE 3.7 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité  $\varphi$ , avec  $e^\beta = 2$ . Pourcentage de censure entre parenthèse.

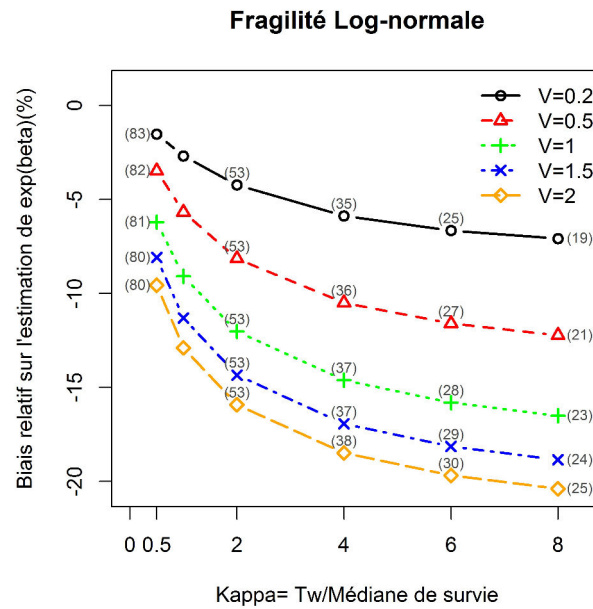


FIGURE 3.8 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité  $V$ , avec  $e^\beta = 2$ . Pourcentage de censure entre parenthèse.



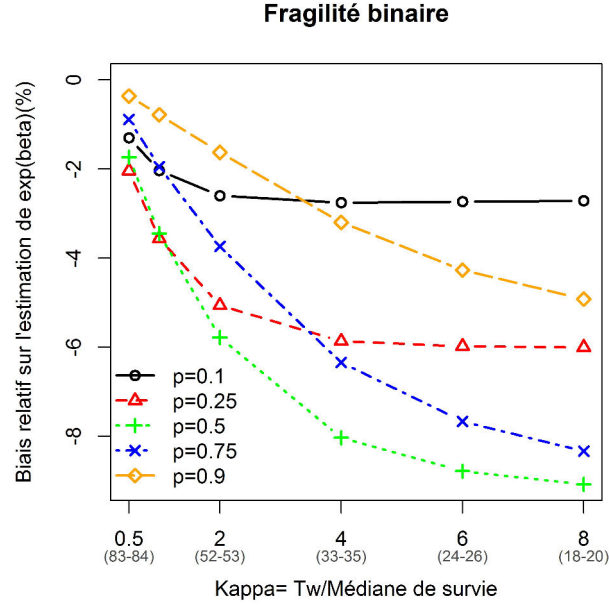


FIGURE 3.9 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et  $p$ , avec  $e^\beta = 2$  et  $\gamma = 1$ . Intervalle des pourcentages correspondant à une valeur de  $\kappa$  donnée entre parenthèse.

courbes obtenues sont proches de celles observées dans le cas où  $T_0 = T_w$ .

Enfin, le biais est calculé avec une censure dépendante du groupe de traitement, à partir de l'équation (2.4). Comme mentionné plus haut, la durée de l'étude est choisie égale à  $T_w$  dans l'un des groupes et  $T_w/k$  dans l'autre, avec  $k = 2$  ou  $4$ . Un tel schéma de censure peut être rencontré dans le cas où des effets indésirables d'un médicament sont attendus dans l'un des groupes de traitement, alors la durée d'étude la plus courte serait planifiée dans ce groupe. Dans ce cas de figure, les temps de censure tendent à être plus court dans le groupe où la durée d'étude est égale à  $T_w/k$ , particulièrement quand  $k$  augmente. Les résultats (Annexe B) montrent que si les temps de censure apparaissent plus tôt dans le groupe expérimental où le risque est plus élevé, les valeurs du biais sont légèrement plus élevées que celles observées dans le cas d'une censure indépendante. A l'inverse, si  $k = 2, 4$  dans l'autre groupe, le biais est plus faible. Cependant, les résultats

restent très proches de ceux obtenus dans le cas indépendant.

Pour illustrer le résultat sur l'influence de la durée de l'étude sur le biais, les résultats d'une étude en oncologie publiés en 2010 ont été repris.

### 3.5 Illustration : l'essai CALGB 9633

La chimiothérapie adjuvante a fait l'objet de plusieurs essais cliniques afin de montrer son bénéfice sur la survie globale chez les patients ayant subi une résection chirurgicale pour un cancer bronchique non à petites cellules (CBNPC). Bien que son efficacité est avérée chez les patients au stade II et IIIA CBNPC, les résultats sont discordants en ce qui concerne les sujets au stade IB (Strauss et al., 2008; Sangha et al., 2010). L'essai CALGB (« Cancer and Leukemia Group B ») 9633 a débuté en 1996 dans le but d'évaluer l'effet d'une chimiothérapie adjuvante (paclitaxel/carboplatin) sur la survie globale dans cette catégorie de sujets. Il inclut 344 patients ayant subi une lobectomie ou une pneumectomie. Ils sont randomisés dans l'un des deux groupes de traitement dans les 4 à 8 semaines suivant la résection : 173 dans le groupe subissant la chimiothérapie adjuvante et 171 dans le groupe témoin. La chimiothérapie consiste en une injection intraveineuse de paclitaxel  $200\text{mg}/\text{m}^2$  durant 3 heures, et de carboplatin avec une aire sous la courbe de  $6\text{mg}/\text{mL}$  par minute pendant 45 à 60 minutes toutes les 3 semaines à répéter 4 fois.

En 2004, avec un suivi médian de 34 mois, les résultats préliminaires ont montré que la chimiothérapie adjuvante permettait une amélioration de la survie, avec un rapport des risques de 0.62 pour un intervalle de confiance de  $[0.44-0.98]$  (Strauss et al., 2004). Cependant, en 2007, avec un suivi médian de 54 mois, la nouvelle analyse ne montrait aucune différence significative entre les deux groupes de traitement. Le rapport des risques obtenu était de 0.80 compris dans l'intervalle de confiance  $[0.60-1.07]$

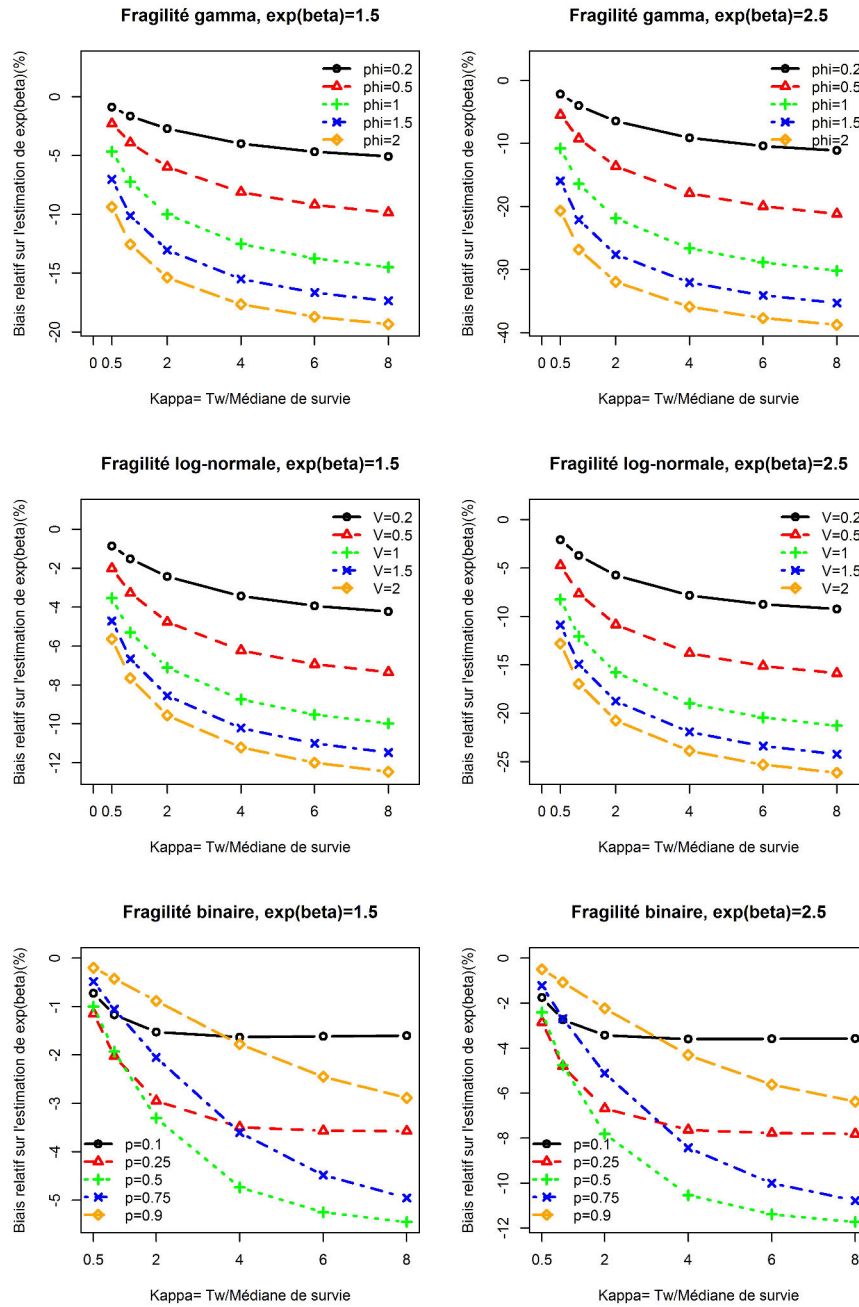


FIGURE 3.10 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité ou du paramètre  $p$ , avec  $e^\beta = 1.5$  ou 2.5.

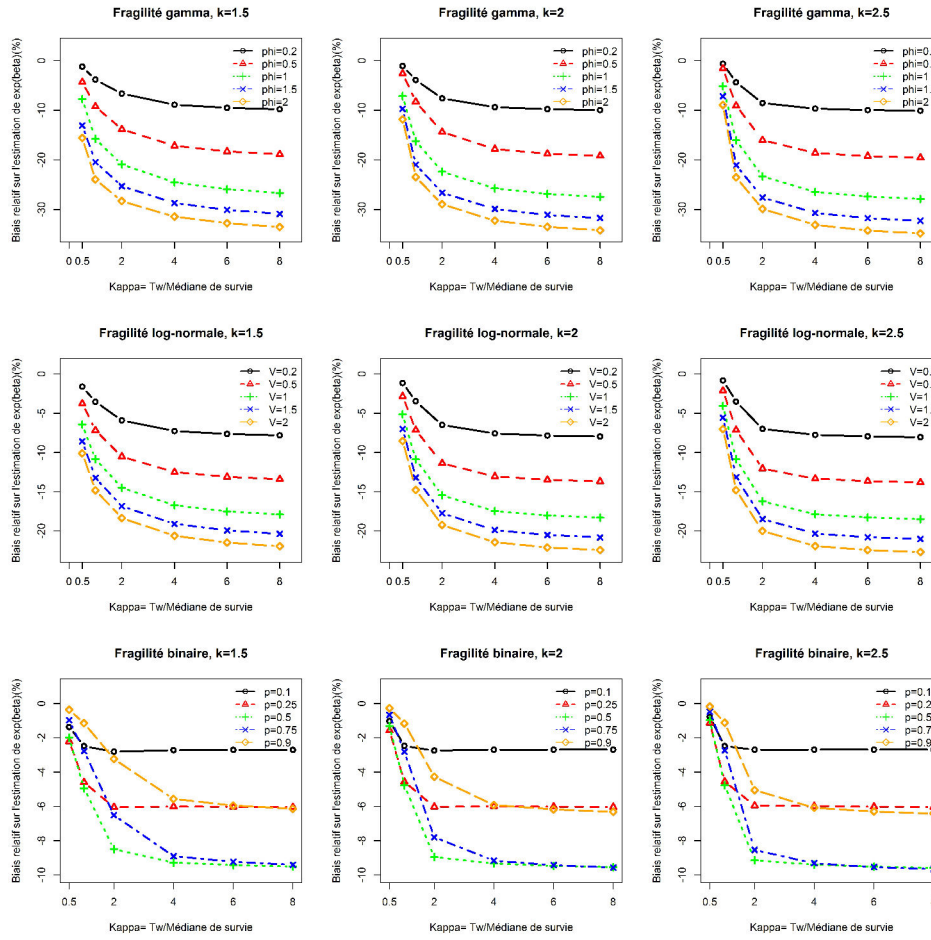


FIGURE 3.11 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité ou du paramètre  $p$ , avec  $e^\beta = 2$ . Le délai de survenue de l'évènement suivant une distribution Weibull( $k,1$ ), avec  $k = 1.5, 2$  ou  $2.5$ .

(Strauss et al., 2008). Cette discordance souleva alors des questions quant à l'utilité de la chimiothérapie pour les malades au stade IB.

Comme il a été récemment discuté par les investigateurs de l'essai CALGB 9633 (Katz et Saad, 2009), quand l'étude était en cours de conception dans les années 90, ils ont décidé de ne pas utiliser les marqueurs cliniques et biologiques bien qu'ils aient réalisé que la population incluse dans l'essai était hétérogène. Par la suite, plusieurs études génomiques sur le CBNPC ont clairement démontré l'hétérogénéité clinique et génomique des adénocarcinomes du poumon au stade IB. Cela soulève la question de savoir si cette hétérogénéité cachée couplée à la durée de l'étude ne serait pas responsable de cette différence observée entre les résultats précoces et tardifs.

Afin d'évaluer si le bénéfice de la chimiothérapie peut effectivement être écarté comme mentionné par quelques oncologistes (Katz et Saad, 2009), nous avons essayé de reproduire le rapport des risques qui a été publié pour l'essai CALGB 9633, en supposant un modèle à risques proportionnels avec une variable fragilité continue unique et indépendante du groupe de traitement. La vraie valeur de l'effet traitement est supposée être égale à 0.60 et le risque de base est choisi comme suivant une distribution Weibull, avec des paramètres d'échelle et de forme respectivement égaux à 0.01 et 3, afin d'obtenir des courbes de survie théoriques (figure 3.12) proches des courbes de Kaplan-Meier observées (Strauss et al., 2008). La fragilité est supposée suivre une distribution gamma avec une espérance et une variance respectivement égales à 1 et 1.5. Une censure de distribution uniforme sur l'intervalle  $(T_w - 7, T_w)$  est considérée afin de représenter l'inclusion des sujets de 1996 à 2004.

En utilisant l'équation 2.4 avec une omission de l'hétérogénéité, l'estimation du maximum de vraisemblance du rapport des risques pour le traitement est de 0.64 avec un suivi de 7 ans et de 0.71 avec un suivi de 11 ans. Dans l'essai, les intervalles de confiance ont été calculés à partir de l'estimateur classique de la variance, c'est-à-dire

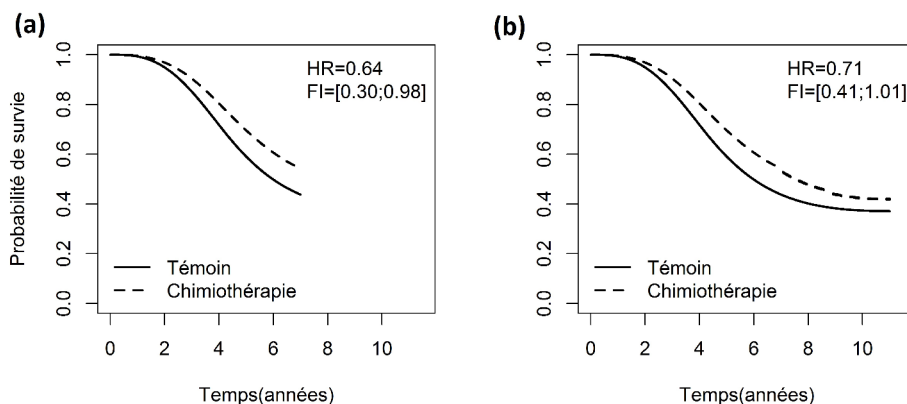


FIGURE 3.12 – Survie théorique pour les groupes témoin et de chimiothérapie avec une durée d'étude de 7ans (a) et 11 ans (b).

l'inverse de l'information de Fisher. Par conséquent, nous avons calculé des intervalles de fluctuation à 90% en utilisant la limite asymptotique de la variance estimée, égale à  $(i(\beta_m^*))^{-1}$  (section 2.2.2), et non la limite de la variance effective de l'estimateur correspondant à l'expression sandwich présentée en section 2.2.2. Ainsi, pour 344 sujets, le premier intervalle est égal à  $[0.30, 0.98]$ , et n'inclut donc pas la valeur 1, contrairement au deuxième intervalle égal à  $[0.41, 1.01]$ .

La figure 3.12 montre les courbes de survie théoriques pour les deux groupes de traitement (témoin et chimiothérapie) avec une durée d'étude de 7 ans (a) et 11 ans (b). Comme attendu et en accord avec les résultats observés dans l'essai CALGB 9633, la différence entre les deux courbes de survie diminue au cours de l'étude. Dans l'ensemble, ces résultats indiquent que l'effet de l'omission de la fragilité peut être assez fort pour expliquer presque complètement l'atténuation observé du rapport des risques au cours du temps, même si les calculs si-dessus ne correspondent pas exactement à la dernière valeur du rapport des risques de 0.80 observée dans l'essai.

### 3.6 Discussion

Le principal résultat de ce chapitre est que la durée de l'étude influence sensiblement le biais asymptotique de l'estimation de l'effet traitement quand la fragilité individuelle n'est pas prise en compte dans le modèle. Aussi, l'impact du type de distribution de la fragilité sur le biais apparaît clairement.

En général, le pourcentage du biais asymptotique obtenu, qui peut atteindre 30%, est du même ordre que celui observé dans la littérature. Par exemple, dans le cas d'une fragilité binaire ou parallèlement d'une covariable binaire dans un modèle de Cox, Schmoor et Schumacher (1997) ont calculé un biais relatif asymptotique  $(\beta_m^* - \beta)/\beta$ , pour une valeur du vrai rapport des risques de  $e^\beta = 2$ . Leurs résultats avec une fragilité pouvant prendre les valeurs 1 ou  $e^\gamma = \log(3)$  et 20% de censure (50% respectivement) peuvent être comparés aux nôtres, où  $e^\beta = 2$ ,  $\gamma = 1$ , et  $\kappa = 8$  ( $\kappa = 2$ , respectivement). Ils trouvent un biais relatif variant entre -5 et -15% (respectivement -3 et -9%). En notant qu'une approximation du premier ordre de notre biais relatif  $\{exp(\beta_m^*) - exp(\beta_m)\} / exp(\beta_m)$  pour  $\beta_m^*$  proche de  $\beta$  est  $(\beta_m^* - \beta)$ , on peut voir sur la figure 3.9 que le biais relatif asymptotique de  $\beta_m^*$  varie entre -4 et -13% (respectivement -2 et -8%).

Nos résultats montrent que dans les situations les plus probables, le biais dépend principalement de la durée de l'étude, qui est fixée lors de la planification de l'essai, et des caractéristiques de la distribution de la fragilité et par conséquent de la (des) covariable(s) oubliées. Étant donné que la variance de la fragilité reflète à la fois la variance de la covariable oubliée et son coefficient de régression, son rôle sur le biais comme observé ici confirme celui du coefficient de régression de la covariable oubliée, qui a été trouvé dans les travaux précédents (Schumacher et al., 1987; Schmoor et Schumacher, 1997). L'influence observée de la vraie valeur de l'effet traitement sur le biais concorde également avec les résultats de la littérature (Gail et al., 1984; Chastang

et al., 1988). Un autre point important est que le biais est sensiblement plus élevé avec des fragilités continues comparativement à une fragilité binaire, autrement dit avec des covariables oubliées continues versus binaire, ce qui ne semble pas avoir été mentionné jusque là. L'importance pratique de ce point devrait être soulignée car dans la plupart des cas, il est raisonnable de penser qu'oublier une unique covariable binaire n'est pas réaliste. Dans plusieurs situations, il est plus probable que plusieurs covariables soient omises, leur combinaison linéaire n'étant pas incluse dans le modèle peut donc être considérée comme approximativement normale. Dans de telles situations, nos résultats montrent que le biais relatif est assez élevé, atteignant 18% quand le risque relatif entre les groupes de traitement est égal à 2 (figure 3.8).

Rappelons que le biais asymptotique à un temps donné est une moyenne pondérée des différences entre les rapports des risques marginaux et la vraie valeur du risque relatif prises à chaque temps précédant ce temps considéré. Ainsi, on peut voir que les variations du BR en fonction du paramètre  $\kappa$  sont en adéquation avec les variations du rapport des risques marginaux observées à la section 3.3. En particulier, quand  $|\gamma| > 1.5$  dans le cas binaire, l'évolution du BR en valeur absolue en fonction de la durée de l'étude n'est pas monotone quelque soit la valeur de  $p$  (Annexe C). En effet, quand  $p \leq 0.5$ , le BR augmente au début de l'étude, atteint un maximum puis décroît. Cette baisse apparaît plus précocement pour les valeurs de  $p$  les plus faibles et n'apparaît pas pour  $p > 0.5$ . De plus, elle apparaît uniquement pour les plus grandes valeurs de  $\gamma$ , et plus  $\gamma$  est élevé plus la baisse est précoce. Ce comportement du BR coïncide avec celui du rapport des risques marginaux. Il peut cependant être souligné que des valeurs de  $\gamma$  supérieures à 1.5 sont plutôt élevées et ont peu de chance d'être rencontrées en pratique.

Le rôle de la proportion de censure, pour laquelle il a été montré qu'elle entraînait une diminution du biais (Schmoor & Schumacher, 1997; Henderson & Oman, 1999; Abrahamowicz et al., 2004), n'est que partiellement confirmée par ces résultats. En



effet, dans le schéma d'étude considéré ici, l'effet de la censure ne reflète pas seulement celui de la durée de l'étude mais également celui de la variance de la fragilité. Cela est visible sur les figures 3.7 et 3.8 qui montrent que le biais augmente non seulement lorsque la proportion de censure liée à la durée de l'étude diminue, mais aussi lorsque la proportion de censure liée à la variance de la fragilité augmente. Par conséquent, le pourcentage de censure ne peut pas être considéré comme un facteur du biais, étant donné que son effet est expliqué par la durée de l'étude couplée à la variance de la fragilité. D'autre part, on peut noter que la distribution de censure a peu d'effet sur le biais. Des distributions de censure exponentielle tronquée et beta ont été explorées (non présenté) et ont abouti à des résultats similaires à ceux obtenus avec une distribution de censure uniforme.

L'exemple illustratif montre que l'hétérogénéité clinique peut expliquer les résultats discordants observés dans l'essai CALB9633. De plus, il apporte des arguments pour ne pas écarter une analyse en sous-groupes qui montrait une amélioration significative de la survie pour les patients ayant de grosses tumeurs (supérieures à 4cm de diamètre) (Strauss et al., 2008). Dans le cadre des essais de prévention à l'infection VIH, l'effet traitement estimé a également été trouvé comme diminuant avec le temps dans des conditions particulières telles qu'une hétérogénéité catégorielle et dans un cadre paramétrique étudié par Auvert et al. (2011). Ces résultats sont confirmés et étendus par ceux présentés dans cette étude. Toujours dans le cadre des essais VIH, le chapitre suivant étudie le biais dans un cas particulier de fragilité dépendante du temps pouvant être rencontrée dans ce contexte.

# Chapitre 4

## Un cas particulier de fragilité dépendante du temps

### 4.1 Motivations

Dans le cadre des maladies infectieuses et particulièrement dans le cadre de l'infection VIH, plusieurs facteurs sont dépendants du temps et le statut « à risque » d'un sujet peut lui-même être intermittent. Les effets de l'omission d'une telle hétérogénéité dépendante du temps sur l'estimation de l'effet traitement sont l'objet de ce chapitre.

Dans les essais de prévention à l'infection VIH, on peut considérer que tout sujet non consommateur de drogue n'est exposé au risque d'infection pendant une période donnée que s'il a au moins un partenaire sexuel infecté durant cette période. Or, dans la majorité des études, le statut VIH du partenaire sexuel est soit indisponible soit imprécis, il n'est alors pas considéré dans l'analyse (IOM (Institute of Medicine), 2008). Ainsi, chaque sujet n'est en fait à risque que sur des intervalles de temps inconnus et les facteurs de risque existant durant ces intervalles, tels que la charge virale du partenaire par exemple, sont inconnus également. Cette situation se rencontre en particulier dans

les populations à haut risque d'infection et instables telles que les prostituées ou les jeunes en Afrique sub-Saharienne.

Le but de ce chapitre est d'évaluer les effets d'une telle hétérogénéité sur l'estimation du rapport des risques. Dans le cas de covariables oubliées constantes, il a été montré que le biais était causé par un déséquilibre du nombre de sujets à risque dans les deux groupes de traitement qui augmentait avec le temps. On peut alors penser que la succession de périodes « à risque » et « non à risque » pour chaque sujet pourrait rétablir l'équilibre au cours du temps et limiter le biais. Cette question sera abordée ici par une étude de simulation. Le modèle sur lequel elle repose utilise une variable fragilité dépendant du temps qui permet de tenir compte à la fois des covariables oubliées et d'un statut « à risque » intermittent.

## 4.2 Modèle de simulations

### 4.2.1 Modèle et notations

En chaque temps  $t$ , chaque sujet  $i$  peut être ou non exposé au risque et les temps aléatoires auxquels peuvent se produire ces changements sont notés  $U_{i1} < U_{i2} < U_{i3} \dots$ . Soient  $V_{ij} = U_{ij} - U_{i(j-1)}$  le temps écoulé entre deux changements potentiels consécutifs,  $u_{i0} = 0$ , et  $w_{ij}$  la fragilité de Bernoulli qui indique si le sujet  $i$  est ou non à risque dans l'intervalle  $V_{ij}$ . Le modèle spécifique au sujet  $i$ , semblable au modèle (2.1), est conditionnel aux  $U_{ij}$  et aux  $w_{ij}$  et le risque instantané du sujet  $i$  sur l'intervalle  $V_{ij}$  est donné par :

$$\lambda(t|w_{ij}, x_i) = \lambda_0(t)w_{ij}e^{\beta x_i}, \quad U_{i(j-1)} \leq t < U_{ij}. \quad (4.1)$$

Le modèle de travail reste le modèle (2.3) à risques proportionnels ne prenant pas

en compte la fragilité.

### 4.2.2 Génération des données

On s'intéresse au biais sur l'estimation de l'effet traitement, en valeur absolue, en fonction de l'espérance et du coefficient de variation de  $V$ . Les données générées reflètent celles d'un essai randomisé comparant deux groupes de traitement selon le modèle (4.1).

La durée de l'essai est notée  $T_w$ , et pour simplification, la durée d'inclusion des sujets est supposée égale à  $T_w$ , de sorte que le temps de censure  $C_i$  de chaque sujet  $i$  est uniformément distribué sur  $[0, T_w]$ . Pour chaque sujet  $i$ , et dans un premier temps, la succession des intervalles est mise en place par tirages aléatoires puis, conditionnellement à celle-ci, le délai de survie est généré. Les intervalles  $V_{ij}$ ,  $j = 1, 2, \dots$  sont générés indépendamment selon une distribution de Weibull. Les temps de changement  $U_{ij}$  potentiels de la fragilité sont déduits des intervalles  $V_{ij}$ , avec  $U_{ij} < C_i$ . Les variables de fragilité  $w_{ij}$ ,  $j = 1, 2, \dots$  proviennent de tirages indépendants de variables de Bernoulli, avec  $p = Pr(w_{ij} = 1)$ .

La distribution du délai de survie  $T_i$  est telle que sur chaque intervalle  $V_{ij}$  où  $w_{ij} = 1$ , le risque instantané est supposé constant et égal à  $\{\lambda_0 \exp(\beta x_i)\}$  où  $x_i$  est la variable indicatrice de l'un des deux groupes de traitement. Sur les intervalles  $V_{ij}$  où  $w_{ij} = 0$ , le risque est supposé nul. Il s'ensuit que le risque cumulé en chaque temps  $t$  tel que  $U_{i(j-1)} \leq t < U_{ij}$  est donné par :

$$\Lambda_i(t) = \Lambda_i(U_{i(j-1)}) + \lambda_0(t - U_{i(j-1)})$$

où la valeur du risque cumulé  $\Lambda_i(U_{i(j-1)})$  dépend du nombre et de l'emplacement des intervalles à risque nul qui précèdent  $U_{i(j-1)}$ . La fonction de survie s'en déduit :

$$S_i(t|w_{ij}, U_{i1}, U_{i2}, \dots, U_{i(j-1)}, x_i) = S_i(U_{i(j-1)}) \exp\{-e^{\beta x_i} \lambda_0 (t - U_{i(j-1)})\}, \quad U_{i(j-1)} \leq t < U_{ij}.$$

Cette fonction est constante sur les intervalles où le risque est nul, et de forme exponentielle sur les autres intervalles. La distribution correspondante diffère des distributions usuelles dans le sens où la probabilité de réalisation de la variable  $T_i$  sur l'intervalle  $[0, +\infty]$  est inférieure à un (Sy et Taylor, 2000). Pour générer le délai de survie, un tirage d'une variable  $z_i$  uniforme sur  $[0, 1]$  est d'abord effectué, puis le temps de survie en est déduit en appliquant à  $z_i$  l'inverse de la fonction ci-dessus au sens suivant. Si la valeur  $S_i^{-1}(z_i)$  appartient à un intervalle de temps où le risque est non nul, cette valeur est celle du temps de survie du sujet. Le cas où la valeur  $z_i$  est égale à une valeur « plateau » de la fonction de survie  $S_i(t)$ , correspondant à un intervalle  $V_{ij}$  de risque nul, a une probabilité nulle de survenir. Cependant, dans ce cas, on peut convenir de choisir  $U_{i(j-1)}$  pour valeur du temps de survie. Finalement, le délai de surveillance du sujet est pris égal à  $\min(T_i, C_i)$ .

Les paramètres de la distribution de Weibull de l'intervalle de temps  $V$  sont choisis tels que l'espérance de  $V$  soit égale à 0.05, 0.25, 0.5, 0.75, 1, 1.5 ou 2, et le coefficient de variation (CV) de  $V$  soit égal à 0.05, 1, 2 ou 4 (figure 4.1). Dans une distribution de Weibull, l'espérance et la variance sont liées entre elles étant donné qu'elles sont toutes les deux dépendantes des deux paramètres de la distribution, c'est-à-dire les paramètres de forme et d'échelle. A l'inverse, le coefficient de variation ne dépend que du paramètre de forme. Par conséquent, chaque membre d'une famille de distribution Weibull qui est caractérisée par un paramètre de forme commun, peut être identifié par son espérance. Dans la suite, les valeurs numériques du biais sont présentées en fonction de l'espérance pour chaque valeur du coefficient de variation. Le risque de base  $\lambda_0$  est fixé à 1 et le

risque relatif entre les deux groupes de traitement vaut 1.5, 2 ou 2.5. La durée de l'étude  $T_w$  est fixée à 2. Nous avons généré 1000 jeux de données en incluant 500 sujets par groupe de traitement.

### 4.3 Résultats

La figure 4.2 montre l'évolution du biais relatif de l'estimation de l'effet traitement en fonction du CV et de l'espérance de  $V$  avec le paramètre  $p$  égal à 0.1 et 0.5. Le nombre moyen de changements de fragilité (NMC) par sujet durant l'étude a été calculé pour chaque CV et chaque espérance de  $V$ . Un intervalle de valeur, correspondant à une espérance donnée, est indiqué entre parenthèses sur les graphiques, la valeur minimale étant atteinte pour  $CV=0.05$  et la maximale pour  $CV=4$ . Le NMC correspond aux changements observables, c'est à dire que la valeur de la fragilité générée pour l'intervalle  $j$  est différente de celle générée à l'intervalle  $j - 1$ . Cette notion diffère donc du nombre de changements potentiels, correspondant au nombre de régénérations de la fragilité. Bien que ces deux notions soient étroitement liées, seule la première est présentée étant donné que seul un changement de la valeur de la fragilité implique un changement de risque du sujet.

Les résultats montrent que pour toutes les valeurs de  $p$  observées, le BR augmente en valeur absolue avec l'espérance de l'intervalle de temps  $V$ . Aussi, quand  $E(V)$  est proche de zéro, le biais augmente avec le coefficient de variation de  $V$ . Quand  $E(V)$  tend vers 2, correspondant à la durée de l'étude  $T_w$ , cette relation est inversée. En effet, avec  $p = 0.1$ , pour  $E(V) = 0.05$ , le BR varie de 2.0% pour  $CV = 0.05$  à 6.8% pour  $CV = 4$ . Pour  $E(V) = 2$ , le BR varie de 31.4% pour  $CV = 0.05$  à 24.9% pour  $CV = 4$ . Le biais diminue légèrement quand  $p = 0.5$  mais les tendances restent les mêmes. Pour  $p = 0.5$ , le BR ne dépasse pas 25.2%.

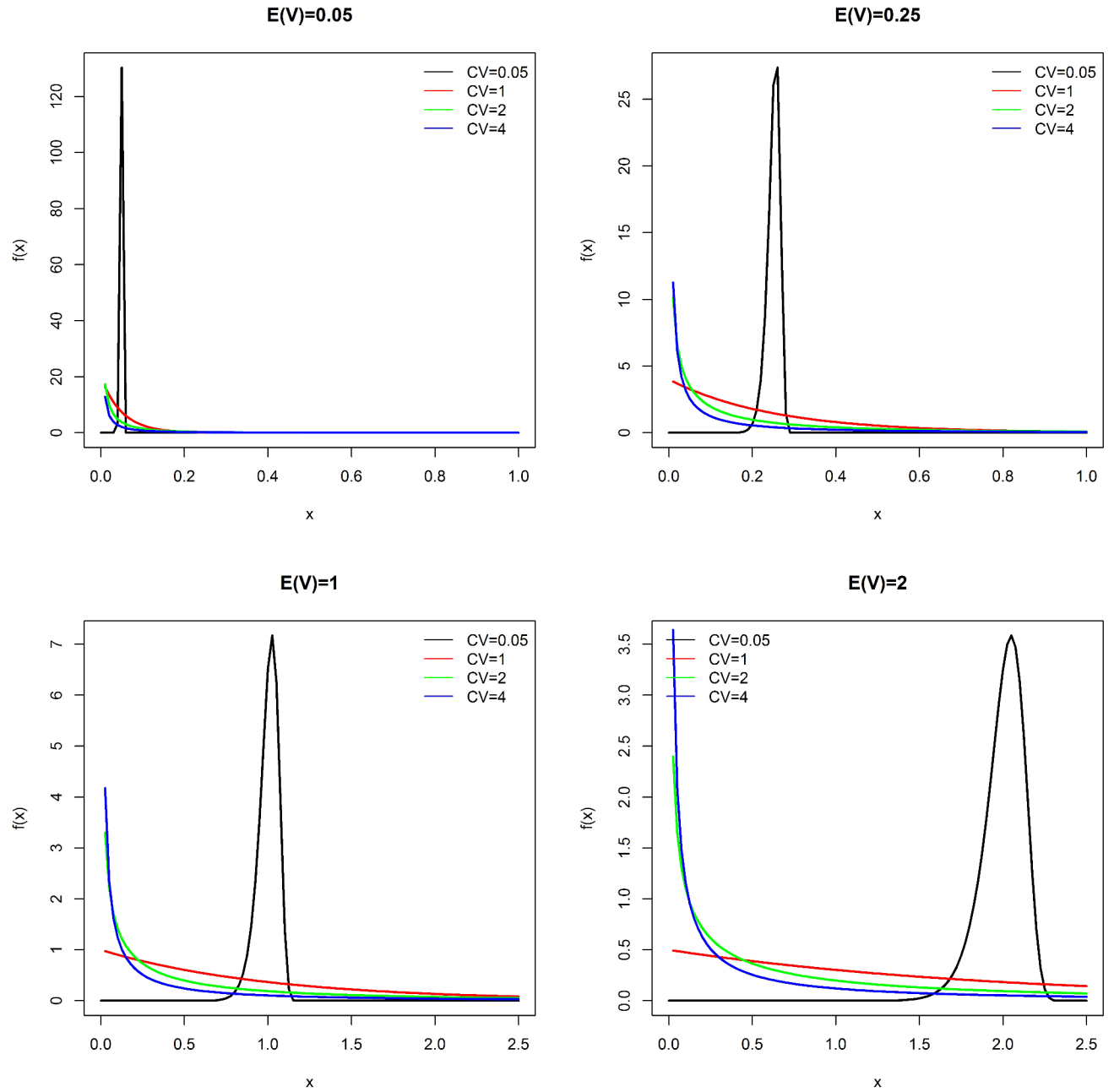


FIGURE 4.1 – Densité de la loi Weibull pour  $CV=0.05, 1, 2$  ou  $4$  et  $E(V)=0.05, 0.25, 1$  ou  $2$ .

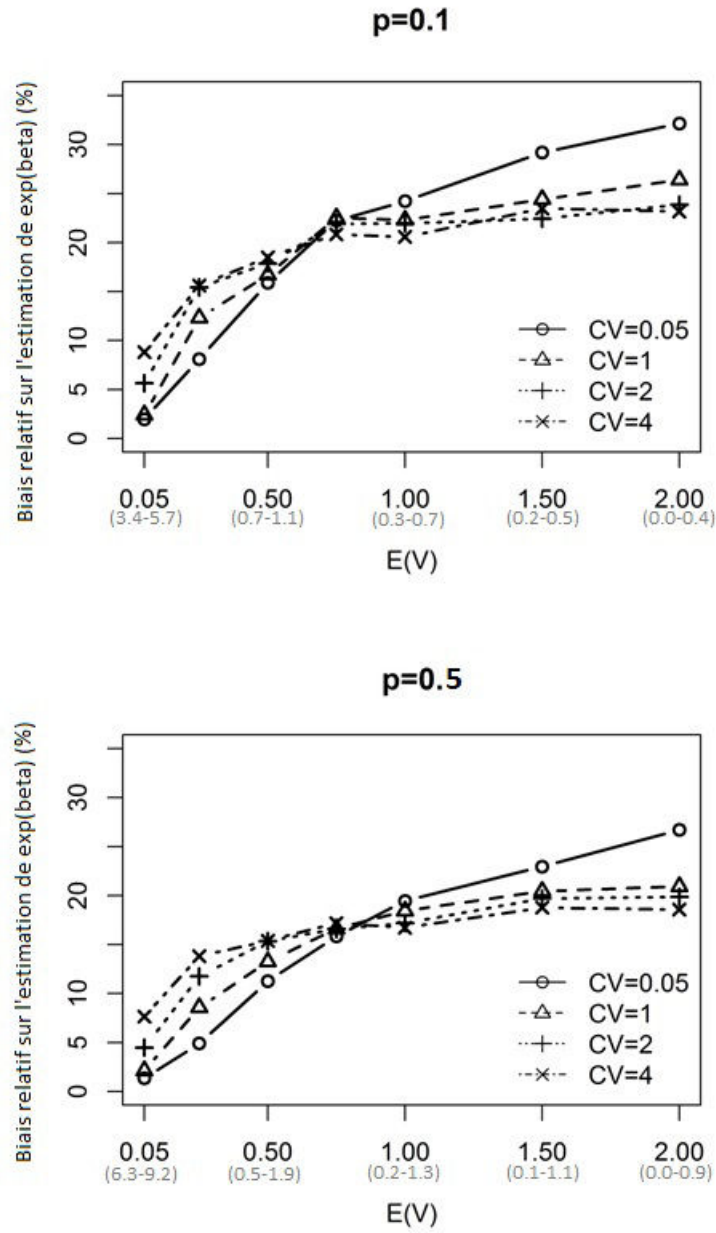


FIGURE 4.2 – Pourcentage du biais relatif de l'estimation de  $e^\beta$  en fonction de  $CV$  et  $E(V)$ , pour  $p = 1$  et  $0.5$ ,  $e^\beta = 2$ ,  $\lambda_0 = 1$  et  $T_w = 2$ . Nombre moyen de changements entre parenthèses.



<b>p</b>	0.1	0.25	0.5	0.75	0.9
<b>RB(%)</b>	32.7	30.4	26.7	18.1	11.2

TABLE 4.1 – Biais relatif sur l'estimation de  $e^\beta$  en fonction de  $p$ , avec une fragilité constante

Le tableau 4.1 regroupe les valeurs du BR quand la fragilité est la même tout le long du suivi, en fonction du paramètre  $p$ . On peut noter que lorsque  $E(V)$  augmente, le BR observé dans le cas d'une fragilité dépendante du temps tend vers la valeur obtenue dans le cas d'une fragilité constante, c'est à dire 32.7% pour  $p = 0.1$  et 26.7% pour  $p = 0.5$ . Ces valeurs sont pratiquement atteintes dans la figure 4.2 avec  $CV = 0.05$  et  $E(V) = 2$ .

Des simulations semblables ont été réalisées avec une corrélation entre les changements de fragilité. Deux valeurs consécutives de la fragilité  $w_{ij}$  et  $w_{i(j-1)}$ , sont liées selon un odds ratio (OR) variant entre 2 et 100. Cela signifie qu'un sujet a plus de chance d'être à risque sur l'intervalle  $j$  s'il était à risque sur l'intervalle  $(j-1)$  que s'il ne l'était pas.

En pratique, la proportion  $P(w_{ij} = 1|w_{i(j-1)} = 1)$  est calculée selon la proportion  $P(w_{ij} = 1|w_{i(j-1)} = 0)$  et l'OR choisis. Par définition :

$$OR = \frac{P(w_{ij} = 1|w_{i(j-1)} = 1)P(w_{ij} = 0|w_{i(j-1)} = 0)}{P(w_{ij} = 1|w_{i(j-1)} = 0)P(w_{ij} = 0|w_{i(j-1)} = 1)}.$$

Par conséquent, la proportion  $P(w_{ij} = 1|w_{i(j-1)} = 1)$  est obtenue par l'égalité suivante :

$$P(w_{ij} = 1|w_{i(j-1)} = 1) = \frac{OR * P(w_{ij} = 1|w_{i(j-1)} = 0)}{1 - P(w_{ij} = 1|w_{i(j-1)} = 0) + OR}.$$

La figure 4.3 montre l'évolution du BR en fonction de l'espérance de  $V$ , avec un OR égal à 10, 50 ou 100, et  $P(w_{ij} = 1|w_{i(j-1)} = 0)$  fixée à 0.25. Les valeurs du biais sont proches de celles observées dans la figure 4.2. Comme dans le cas non-corrélé,

le BR augmente en valeur absolue avec l'espérance de l'intervalle de temps  $V$ . Aussi, l'inversion de l'effet du CV sur le biais est plus précoce quand l'OR augmente : autour de  $E(V) = 0.4$  pour  $OR = 10$  contre  $E(V) = 0.2$  pour  $OR = 100$ . Cependant, qu'importe la valeur de l'OR, le BR tend vers la même limite de 30.4%, valeur cohérente avec la valeur obtenue dans le tableau 4.1 dans le cas d'une fragilité constante. Les mêmes tendances sont observées pour des valeurs de  $P(w_{ij} = 1 | w_{i(j-1)} = 0)$  différentes.

D'autres résultats sont obtenus avec une durée d'étude  $T_w = 4$  (figure 4.4) ou en considérant une distribution de Weibull (figure 4.5) pour la survenue de l'évènement, avec des paramètres de forme et d'échelle respectivement égaux à 1 et 2, avec des changements de fragilité non-corrélés. La même association positive est observée entre le biais et l'espérance de  $V$ .

## 4.4 Discussion

En conclusion, le biais sur l'estimation de l'effet traitement résultant d'une omission de la fragilité individuelle dans le modèle, est moins élevé lorsque la fragilité est redistribuée une ou plusieurs fois au cours du suivi, comparativement au cas où la fragilité est indépendante du temps. Plus l'espérance du délai de changement de fragilité est faible, plus le biais en valeur absolue est faible. Cependant, le biais relatif reste important et ne descend en dessous du seuil de 5% que pour  $E(V) = 0.05$  et  $CV \leq 1$ . Quand  $E(V) \geq 0.5$ , il excède les 10% pour  $p = 0.5$  et les 15% pour  $p = 0.1$ , qu'importe la valeur du coefficient de variation. Comme il a déjà été souligné dans le cas d'une fragilité indépendante du temps, cette étude montre que dans un tel cas de fragilité régénérée au cours du suivi, tel qu'il est rencontré dans les essais de prévention à l'infection VIH, la sous-estimation de l'effet traitement persiste bien qu'atténuée, et ne peut être négligée.

D'autre part, pour une valeur du CV donnée, le NMC diminue lorsque l'espérance

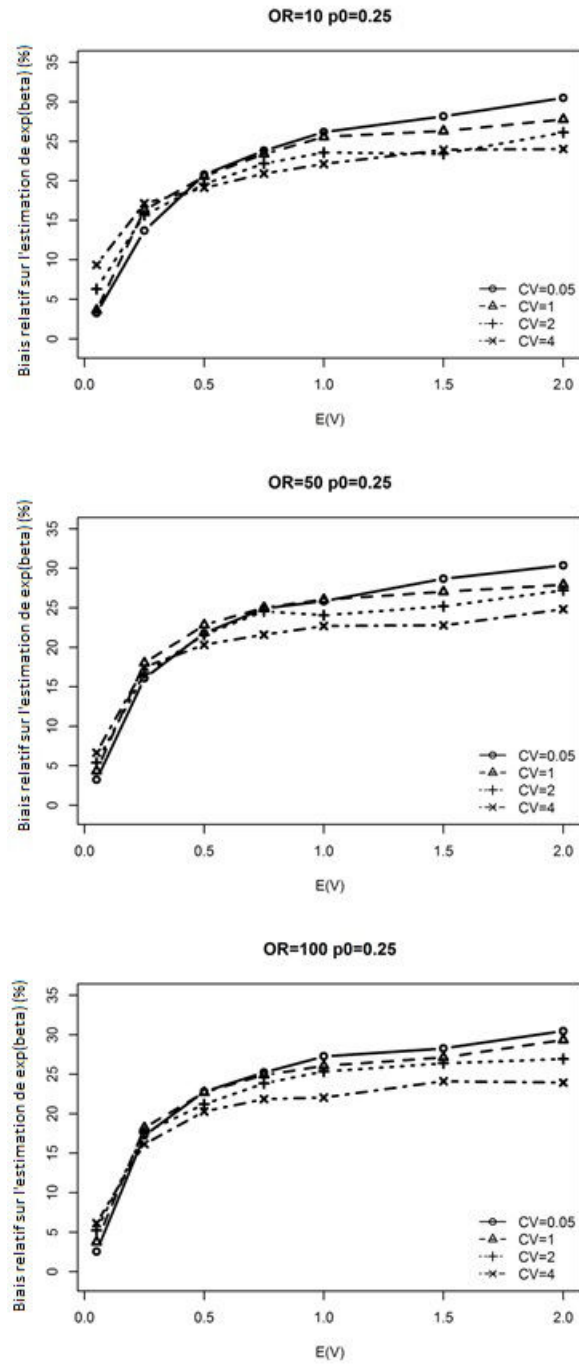


FIGURE 4.3 – Pourcentage du biais relatif de l'estimation de  $e^\beta$  en fonction de  $CV$  et  $E(V)$ , pour  $OR=10, 50$  et  $100$ ,  $e^\beta = 2$ ,  $\lambda_0 = 1$ .

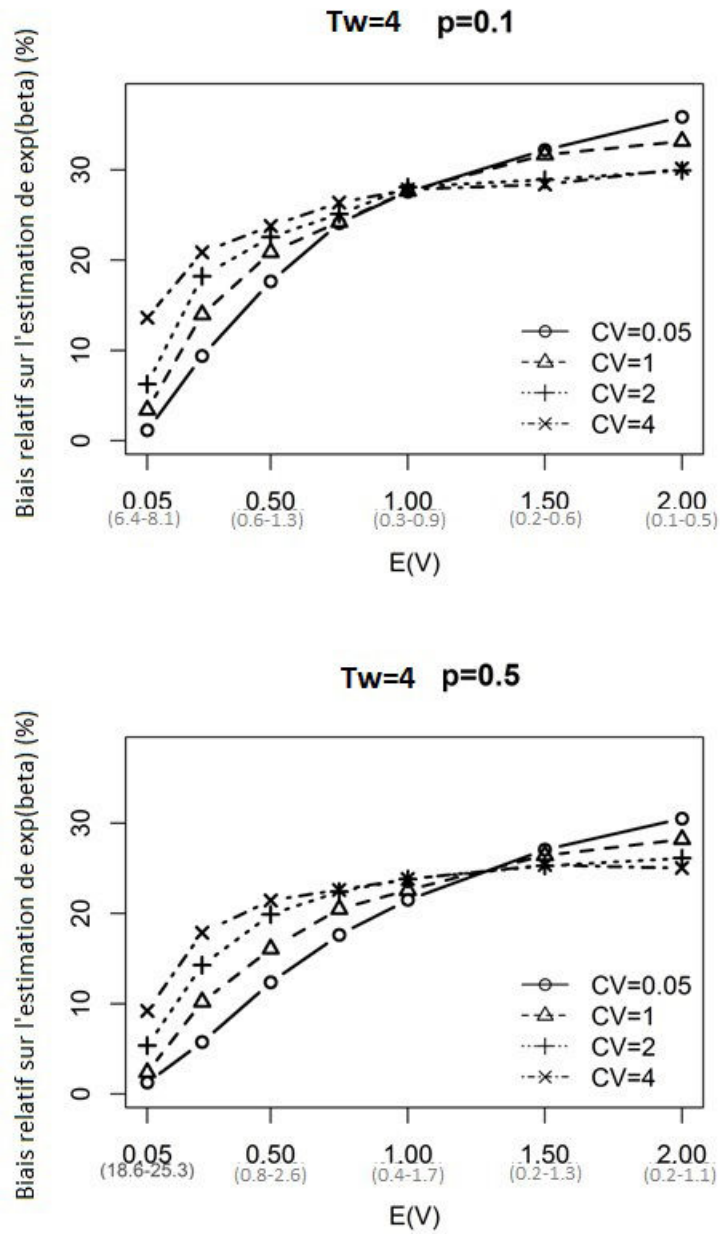


FIGURE 4.4 – Pourcentage du biais relatif de l'estimation de  $e^\beta$  en fonction de  $CV$  et  $E(V)$ , pour  $p = 1$  et  $0.5$ ,  $e^\beta = 2$ ,  $\lambda_0 = 1$  et  $T_w = 4$ . Nombre moyen de changements entre parenthèses.

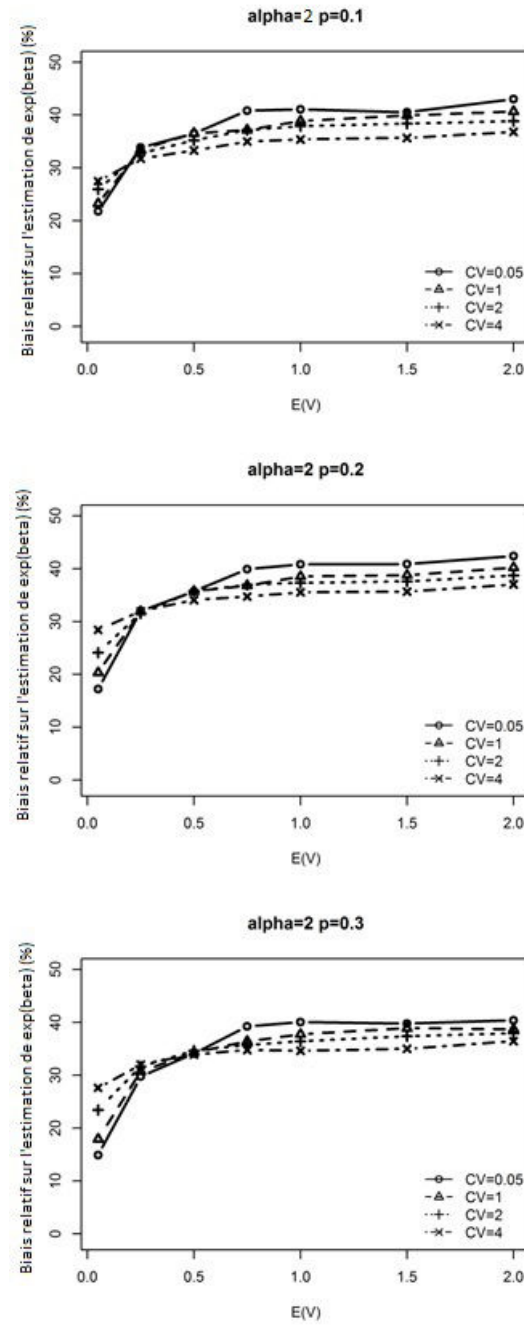


FIGURE 4.5 – Pourcentage du biais relatif de l'estimation de  $e^\beta$  en fonction de  $CV$  et  $E(V)$ , pour modèle de Weibull, paramètres de forme et d'échelle respectivement égaux à 1 et 2,  $p=0.1, 0.2$  ou  $0.3$ .  $e^\beta = 2$ ,  $\lambda_0 = 1$  et  $T_w = 2$ .

du délai de changement  $V$  augmente. Alors, pour une valeur de  $CV$  donnée, le biais augmente en valeur absolue lorsque le nombre de changements diminue. Cependant, pour une valeur de  $E(V)$  donnée, le NMC augmente également avec le coefficient de variation. L'association entre le NMC et le biais est donc positive lorsque  $E(V)$  est proche de 0, et négative lorsque  $E(V)$  est proche de 2. Par conséquent, le nombre de changements ne peut pas être considéré comme un facteur du biais.

Aussi, il est important de souligner que la valeur du biais relatif avec une fragilité dépendante du temps tend vers celle obtenue avec une fragilité constante lorsque l'espérance de  $V$  augmente. Avec un  $CV$  de 0.05, cette limite est quasiment atteinte avec  $E(V) = 2$ . A titre indicatif, avec  $CV = 4$ , une différence inférieure à 0.5% entre le biais relatif et sa limite nécessite une  $E(V)$  supérieure à 100000 (résultat non présenté). Le paramètre d'échelle augmente avec  $E(V)$ , mais moins rapidement si le  $CV$  est faible. Par exemple, avec  $CV = 0.05$ , le paramètre d'échelle varie entre 0.03 pour  $E(V) = 0.05$  et 1.15 pour  $E(V) = 2$ . Avec  $CV = 4$ , il varie entre 0.05 et 2.17, respectivement. Cela peut expliquer la différence de pente pour l'augmentation du biais en fonction de l'espérance de  $V$  selon la valeur du  $CV$ .



## Chapitre 5

# Utilisation d'un schéma en cross-over pour une population hétérogène

Comme il a été précisé précédemment, les avantages de l'utilisation d'un schéma en cross-over pour une population hétérogène ont été explorés par Auvert et al. (2011) dans un cadre paramétrique avec une fragilité en quatre classes, et par Buyze et Goetghebeur (2011) à la fois dans un cadre paramétrique et semi-paramétrique, avec une fragilité gamma ou log-normale. Ces deux études se placent dans le contexte des essais de prévention à l'infection VIH et se basent sur des études de simulation. Dans ce chapitre, la distribution asymptotique du biais présentée au Chapitre 2 permet de préciser par le calcul les résultats obtenus dans ces articles. La comparaison entre les schémas parallèle classique et en cross-over portera particulièrement sur la valeur de ce biais. L'efficacité relative asymptotique des tests de comparaison des deux groupes de traitement auxquels conduisent chacun des deux schémas sera également considérée.



## 5.1 Adaptation des modèles au schéma en cross-over

Le modèle spécifique au sujet définie par (2.1) peut être adapté à un schéma en cross-over en considérant une variable indicatrice du groupe de traitement dépendante du temps  $x_i(t)$ , constante avant et après un temps de changement de groupe  $T_{switch}$  fixé. Ainsi, pour chacun des sujets  $i$ , la fonction de risque est alors définie par :

$$\lambda(t|w_i, x_i(t)) = \lambda_0(t)w_i e^{\beta x_i(t)},$$

la valeur de  $x_i(t)$  pouvant être obtenue sachant le groupe de traitement du sujet  $i$  au début de l'étude  $x_i(t_0)$ , et le temps de changement de groupe  $T_{switch}$  :

$$x_i(t|x_i(t_0) = r, T_{switch}) = \begin{cases} r & \text{si } t \leq T_{switch} \\ (1 - r) & \text{si } t > T_{switch} \end{cases}$$

avec  $r = 0$  ou  $1$ .

Dans le schéma cross-over, le test du Logrank est défini comme le test du score calculé avec la covariable  $x_i(t)$ , dont la statistique est donnée par :

$$\sum_{i=1}^n \int_0^\infty \left( x_i(t) - \frac{n_1(t)}{n_0(t) + n_1(t)} \right) dN_i(t)$$

selon les mêmes notations utilisées au Chapitre 2. Dans la suite de ce chapitre, l'efficacité relative asymptotique (ERA) concerne le test basé sur cette statistique et le Logrank habituel déduit du modèle de travail. Des précisions concernant le calcul de l'ERA sont apportées dans l'Annexe D. Aussi, on suppose une censure fixée à la durée de l'étude pour tous les sujets.

## 5.2 Comparaison avec le schéma parallèle classique

Dans la première sous-section, les schémas parallèle (SP) et cross-over (SC) sont comparés en supposant une fragilité de distribution gamma ou log-normale. Une fragilité en quatre classes, incluant une classe de masse nulle, est étudiée dans la seconde sous-section.

### 5.2.1 Fragilités continues

Les distributions gamma et log-normale représentent les choix les plus courants pour la fragilité dans la littérature, et ont été utilisés plus particulièrement dans l'étude de Buyze et Goetghebeur (2011). Dans les deux cas, l'espérance de la fragilité a été fixée à 1, sa variance pouvant prendre les valeurs 0.5, 1 ou 2. Le délai de survenue de l'évènement suit une distribution exponentielle avec un risque de base égal à 0.3. Le risque relatif entre les deux groupes de traitement prend les valeurs 2, 3 ou 4. La durée de l'étude est fixée à 2. Pour le schéma en cross-over, le suivi est divisé, comme souvent dans les études en cross-over, en deux périodes de même durée, donc égale à 1.

Les tableaux 5.1 comparent les schémas parallèle et cross-over selon plusieurs critères, en utilisant une fragilité gamma ou log-normale, respectivement, pour une variance égale à 1. La limite  $\beta_m^*$  de l'EMV du log de l'effet traitement est calculée pour chaque schéma à partir de l'équation (2.4), le pourcentage du biais relatif (BR) sur l'estimation de  $\beta$  en est déduit. La variance asymptotique  $V^*$  de  $\sqrt{n}(\hat{\beta}_m - \beta_m^*)$  (2.2) et l'efficacité relative asymptotique sont également présentées.

Les résultats montrent que le biais est significativement moins élevé, en valeur absolue, avec un schéma en cross-over comparativement à un schéma parallèle, qu'importe la valeur de  $\beta$ . Par exemple, pour la fragilité gamma et  $e^\beta = 2$ , le BR, d'une valeur de 23.0% en valeur absolue sous le SP, chute à 4.8% sous le SC. Pour la fragilité log-

Fragilité Gamma	$e^\beta = 2$ $\beta \approx 0.693$		$e^\beta = 3$ $\beta \approx 1.098$		$e^\beta = 4$ $\beta \approx 1.386$	
Schéma	SP	SC	SP	SC	SP	SC
$\beta_m^*$	0.534	0.660	0.817	1.020	1.004	1.256
BR(%)	-23.0	-4.8	-25.7	-7.1	-27.6	-9.4
$V^*$	9.0	9.2	8.3	9.0	8.0	9.0
ERA	1.55		1.61		1.64	

Fragilité Log-normale	$e^\beta = 2$ $\beta \approx 0.693$		$e^\beta = 3$ $\beta \approx 1.098$		$e^\beta = 4$ $\beta \approx 1.386$	
Schéma	SP	SC	SP	SC	SP	SC
$\beta_m^*$	0.575	0.665	0.894	1.036	1.113	1.285
BR(%)	-16.9	-4.0	-18.4	-5.7	-19.7	-7.3
$V^*$	8.6	8.9	8.0	8.7	7.7	8.7
ERA	1.36		1.41		1.44	

TABLE 5.1 – Comparaison des biais et variance asymptotiques pour l'estimation de  $\beta$  sous les SC et SP, et efficacité relative asymptotique du test du Logrank pour SC vs SP,  $V(w_i) = 1$ .

normale, le BR passe de 16.9% à 4.0%. Pour les deux distributions de fragilité, le biais connaît une baisse de 60% à 80% avec le SC par rapport au SP.

La variance asymptotique est plus élevée de quelques dixièmes avec le SC. Toutefois, l'ERA est nettement supérieure à 1 dans tous les cas considérés et traduit une meilleure puissance du test de comparaison avec l'utilisation du SC. Pour la fragilité gamma, l'ERA varie entre 1.55 pour  $e^\beta = 2$  et 1.64 pour  $e^\beta = 4$ . Pour la fragilité log-normale, elle varie entre 1.36 pour  $e^\beta = 2$  et 1.44 pour  $e^\beta = 4$ .

Les mêmes tendances sont observées pour les autres valeurs de variance de fragilité considérées.

### 5.2.2 Fragilité catégorielle

Dans cette partie, la fragilité suit une distribution catégorielle en quatre classes, avec une classe de masse nulle, en référence au papier de Auvert et al. (2011), qui ont cherché à se rapprocher de ce qui pouvait être rencontré dans les essais de prévention à

Fragilité catégorielle	$e^\beta = 2$ $\beta \approx 0.693$		$e^\beta = 3$ $\beta \approx 1.098$		$e^\beta = 4$ $\beta \approx 1.386$	
Schéma	SP	SC	SP	SC	SP	SC
$\beta_m^*$	0.483	0.673	0.703	1.047	0.825	1.295
BR(%)	-30.2	-2.9	-36.0	-4.7	-40.5	-6.6
$V^*$	41.4	42.7	37.4	41.5	37.0	41.9
ERA	1.93		2.18		2.40	

TABLE 5.2 – Fragilité de distribution catégorielle : Comparaison des biais et variance asymptotiques pour l'estimation de  $\beta$  sous les SC et SP, et efficacité relative asymptotique du test du Logrank pour SC vs SP.

l'infection VIH. Le vecteur probabilité est noté  $P = (p_1, p_2, p_3, p_4)$  et le vecteur fragilité est égal à  $W = (0, 1, 2, 3)$ . Le risque de base est fixé à 0.2. Les autres paramètres sont fixés comme dans les cas de fragilités continues en 5.2.1.

Le tableau 5.2 montre les résultats pour  $P = (0.8, 0.15, 0.04, 0.01)$ . La classe de masse nulle a donc une probabilité de 80%. Cette valeur peut sembler élevée mais reste une limite basse en termes de risque d'infection VIH en population générale.

Comme dans les cas de fragilités continues, le biais relatif est plus faible sous un SC que sous un SP. Par exemple, pour  $e^\beta = 2$ , le BR chute de 30.2% à 2.9% en valeur absolue. Ainsi, selon la valeur de  $\beta$ , le biais diminue entre 80% et 90% avec un SC comparativement à un SP. La variance asymptotique reste plus élevée avec un SC. Par ailleurs, on peut remarquer que dans les deux schémas, on a des valeurs de variance près de trois fois plus élevées que dans les cas continus. Cependant il faut souligner le fait que l'on a une grande partie de notre population avec un risque nul donc une grande variabilité des sujets quant au risque de survenue de l'évènement. Le bénéfice en termes d'efficacité du test de comparaison est plus accentué que dans les cas de fragilités continues précédents. En effet, l'ERA varie entre 1.93 pour  $e^\beta = 2$  et 2.40 pour  $e^\beta = 4$ . Les résultats sont similaires avec d'autres valeurs du vecteur  $P$ , avec notamment un biais plus faible pour les deux schémas lorsque la valeur de  $p_1$ , la proportion de sujets ayant un risque nul, est plus faible.

Bien que les avantages de l'utilisation d'un schéma en cross-over soient évidents, ces résultats sont obtenus avec un changement de groupe fixé à la moitié du suivi. On peut se demander s'il ne serait pas possible d'améliorer encore les résultats en faisant varier la valeur de  $T_{switch}$ .

### 5.3 Étude du temps de changement de groupe optimal

D'après les résultats de la section précédente, on voit que le gain de puissance apporté par l'utilisation d'un schéma en cross-over est en grande partie basé sur la diminution du biais étant donné que la variance varie peu d'un schéma d'étude à l'autre. Le biais constitue donc le critère choisi pour définir le temps de changement de groupe optimal, noté  $T_{switch}^0$ , conduisant à une erreur d'estimation minimale. Les facteurs qui pourraient influencer sur celui-ci sont aussi explorés.

Dans un premier temps, les figures 5.1, 5.2 et 5.3 représentent les variations du biais relatif asymptotique pour l'estimation de  $\beta$  en fonction du temps de changement de groupe  $T_{switch}$ , en faisant varier les paramètres de distribution de la fragilité, ainsi que la valeur de  $\beta$ .

Les valeurs choisies pour  $exp(\beta)$  sont les mêmes que dans la section précédente. Les figures 5.1 et 5.2 présentent les cas de distributions de fragilité gamma et log-normale, respectivement, avec une espérance de 1 et une variance égale à 0.5, 1 ou 2. La figure 5.3 présente le cas de la distribution catégorielle. En plus des résultats pour un vecteur  $P = (0.80, 0.15, 0.04, 0.01)$  comme utilisé précédemment, ceux pour  $P = (0.50, 0.30, 0.15, 0.05)$  sont aussi présentés.

Les résultats montrent que la valeur du biais peut être réduite, en termes de valeur absolue, en modifiant la valeur du temps de changement de groupe de traitement qui, pour tous les cas étudiés, se situe avant la moitié du suivi. Dans le cas des distributions

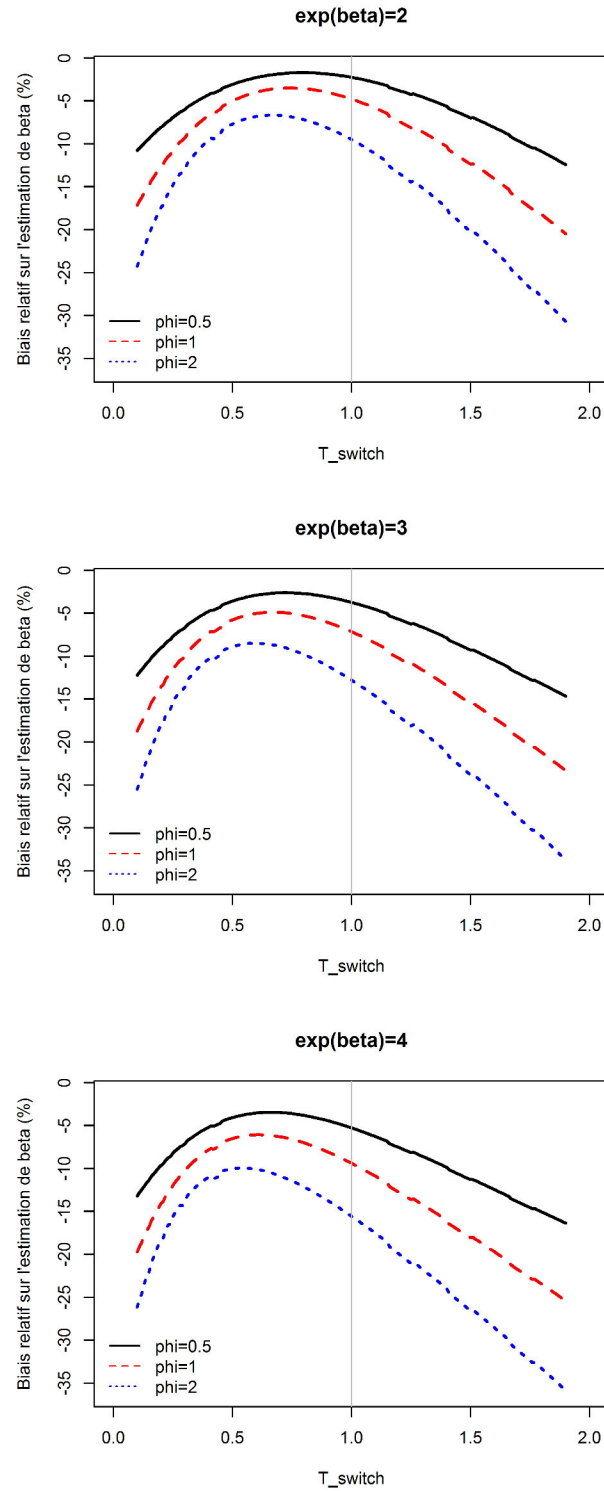


FIGURE 5.1 – Fragilité de distribution gamma : Pourcentage du biais relatif asymptotique sur l'estimation de  $\beta$  en fonction du temps de changement de groupes  $T_{switch}$ , avec  $T_w = 2$ .

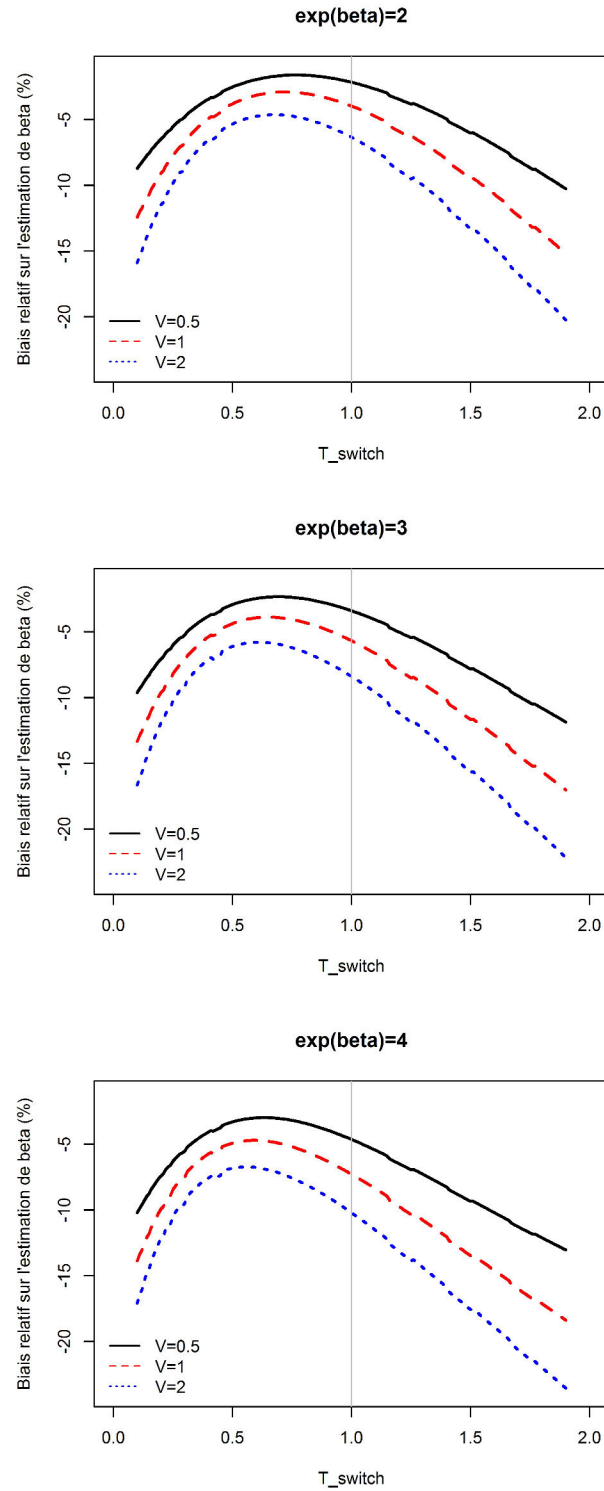


FIGURE 5.2 – Fragilité de distribution log-normale : Pourcentage du biais relatif asymptotique sur l'estimation de  $\beta$  en fonction du temps de changement de groupes  $T_{switch}$ , avec  $T_w = 2$ .

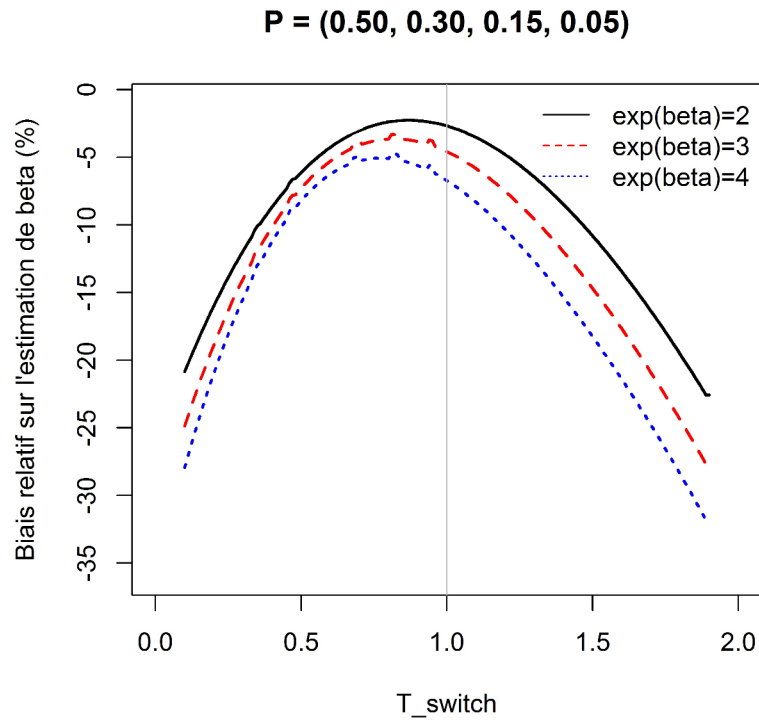
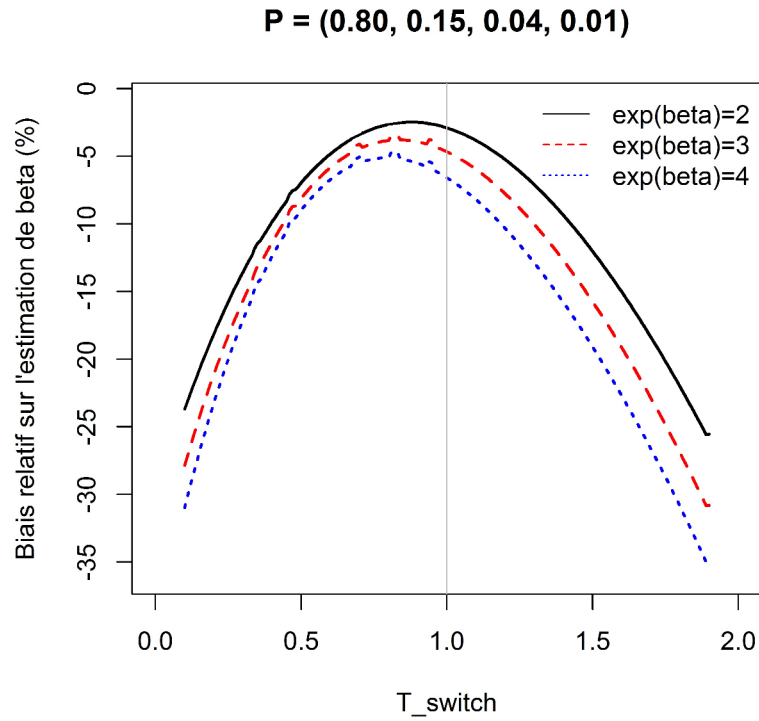


FIGURE 5.3 – Fragilité de distribution catégorielle : Pourcentage du biais relatif asymptotique sur l'estimation de  $\beta$  en fonction du temps de changement de groupes  $T_{\text{switch}}$ , avec  $T_w = 2$ .



de fragilité continues, un changement à  $T_{switch}^0$  permet de frôler la barre des 90% de baisse du biais relatif. Par exemple, pour  $exp(\beta) = 2$  et une fragilité gamma de variance 1, on passe de 23% de biais avec un SP, observé au tableau 5.1, à 2.4% avec un SC programmant un changement de groupe à un temps optimal, contre 4.8% pour un changement à  $T_{switch} = 1$ . Pour la fragilité catégorielle,  $T_{switch}^0$  est plus proche de 1 donc la valeur minimale du biais est très proche de celle obtenue pour un suivi divisé en deux périodes égales. Par exemple, pour  $exp(\beta) = 2$  et un vecteur de probabilité  $P = (0.80, 0.15, 0.04, 0.01)$ , le biais minimal est de 2.6% alors qu'on obtient déjà un biais de 2.9% avec un changement de groupe à 1 (tableau 5.2).

Afin de définir plus précisément les facteurs pouvant influencer sur la valeur du temps de changement  $T_{switch}^0$ , celle-ci a été représentée en fonction de la durée de l'essai  $T_w$ , de la variance de la fragilité pour les distributions continues, ou du vecteur  $P$  pour la distribution catégorielle, pour les valeurs de l'effet traitement considérées précédemment. Les résultats obtenus sont présentés dans les figures 5.4, 5.5 et 5.6.

Tout d'abord, on peut observer que dans tous les cas considérés, la valeur du temps de changement optimal  $T_{switch}^0$  augmente avec la durée de l'essai  $T_w$ . Toutefois, il est important de noter que l'augmentation de  $T_{switch}^0$  n'est pas proportionnelle à celle de  $T_w$  et que le changement optimal s'éloigne du temps correspondant à la moitié du suivi lorsque  $T_w$  augmente. Par exemple, pour une fragilité gamma et  $exp(\beta) = 2$ , le rapport  $T_{switch}^0/T_w$  se situe entre 0.3 et 0.4 pour  $T_w = 2$  selon la variance de la fragilité, alors qu'il est plutôt entre 0.15 et 0.2 pour  $T_w = 10$ . Aussi, dans le cas d'une fragilité gamma, on peut noter que, pour une durée de l'essai  $T_w$  fixée, le temps  $T_{switch}^0$  augmente lorsque la variance de la fragilité diminue. Cet effet est moindre pour la distribution log-normale pour laquelle les courbes restent très proches. D'ailleurs, une inversion de l'effet de la variance est même observée pour  $T_w > 8$  et  $e^\beta \geq 2$ . Les valeurs de  $T_{switch}^0$  diffèrent peu également pour les deux cas catégoriels considérés, étant tout de même légèrement

moins élevées pour le vecteur  $P$  comprenant la plus petite valeur de  $p_1$ . D'autre part, pour une valeur de  $T_w$  fixée, une diminution de la valeur de  $T_{switch}^0$  est observée lorsque la valeur de  $\beta$  augmente.

## 5.4 Discussion

Ce chapitre montre que l'utilisation d'un schéma en cross-over entraîne une importante diminution du biais sur l'estimation du coefficient  $\beta$  par rapport à un schéma parallèle classique. Avec un changement de groupe de traitement fixé à la moitié de la durée de l'étude, on obtient une baisse entre 60% et 80% pour les cas de fragilités continues, et jusqu'à 90% si le temps de changement de groupe est choisi de façon optimal. On observe également une baisse du biais relatif pouvant dépasser les 90% dans le cas catégoriel considéré. La variance asymptotique de l'EMV est constatée légèrement plus élevée avec un SC comparativement au SP. Toutefois, cela n'empêche pas un net gain de puissance du test du Logrank avec l'utilisation du cross-over dans tous les cas considérés, avec une efficacité asymptotique pouvant notamment passer du simple au double dans le cas catégoriel. Tous ces résultats sont cohérents avec ceux obtenus dans les deux principaux papiers cités. A titre indicatif, les calculs ont été refaits pour  $\exp(\beta) = 1.5$  ( $\beta \approx 0.405$ ) et  $T_w = 10$  afin de les comparer avec ceux obtenus par simulations par Buyze et Goetghebeur (2011), et des valeurs très proches ont été retrouvées. En effet, nous avons obtenu des valeurs de  $\beta_m^*$  égales à 0.246 pour un SP et 0.339 pour un SC, alors que Buyze et Goetghebeur ont présenté des valeurs égales à 0.248 et 0.319 respectivement. Aussi, les valeurs de l'efficacité relative asymptotique observées sont du même ordre que dans notre étude.

D'autre part, deux facteurs principaux ont été identifiés comme influant sur le temps de changement de groupe optimal  $T_{switch}^0$  ; une association négative a été observée entre

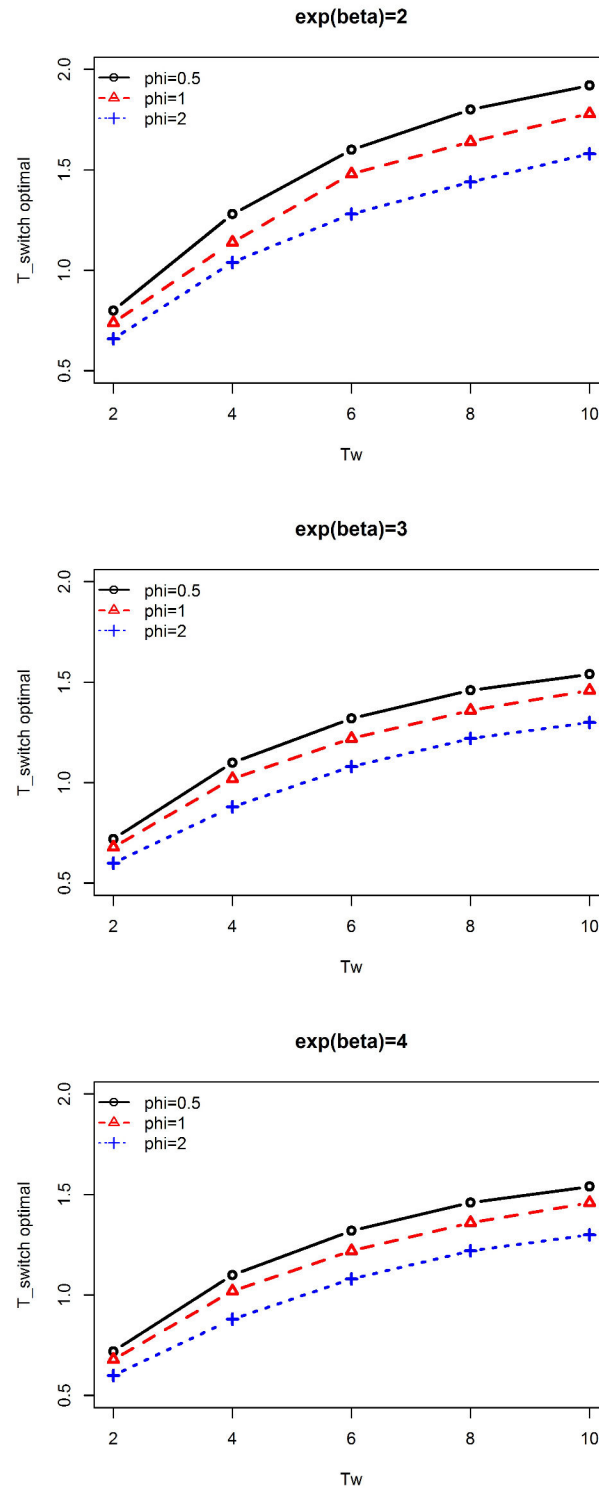


FIGURE 5.4 – Fragilité de distribution log-normale :  $T_{switch}^0$  en fonction de  $T_w$  et de  $\varphi$  pour  $e^\beta = 2, 3$  ou  $4$ .

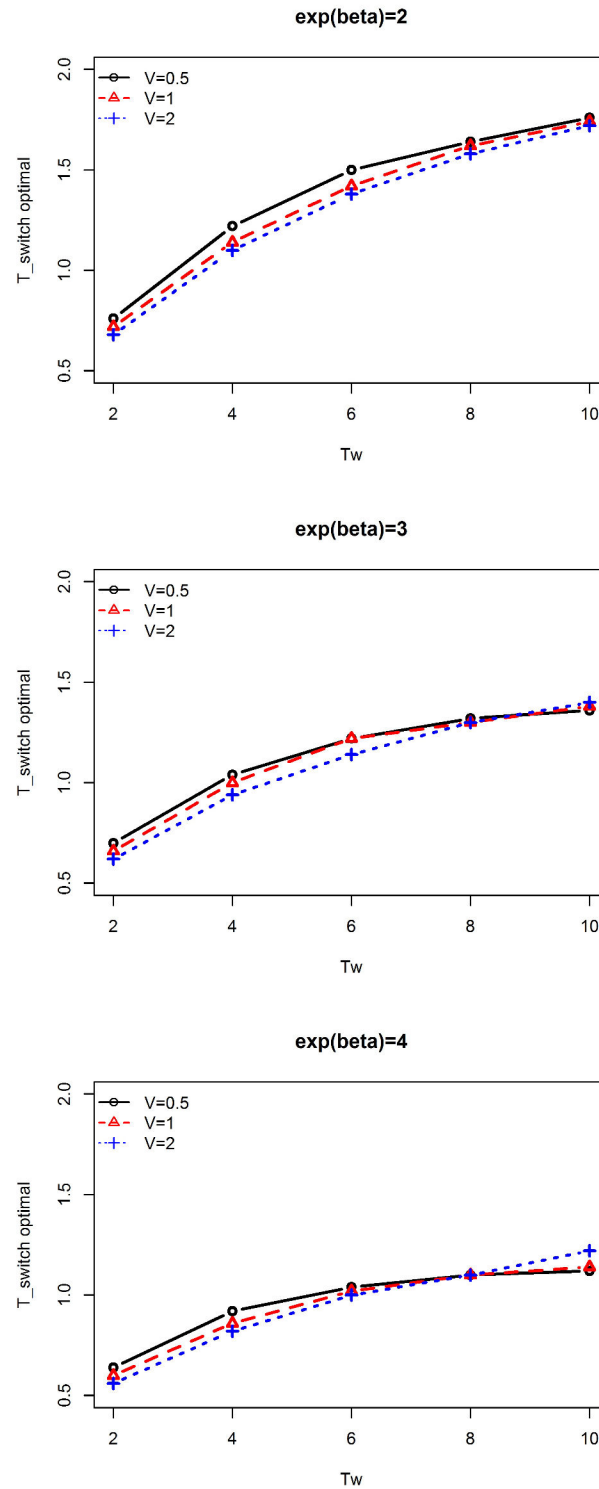


FIGURE 5.5 – Fragilité de distribution log-normale :  $T_{switch}^0$  en fonction de  $T_w$  et de  $V$  pour  $e^\beta = 2, 3$  ou  $4$ .

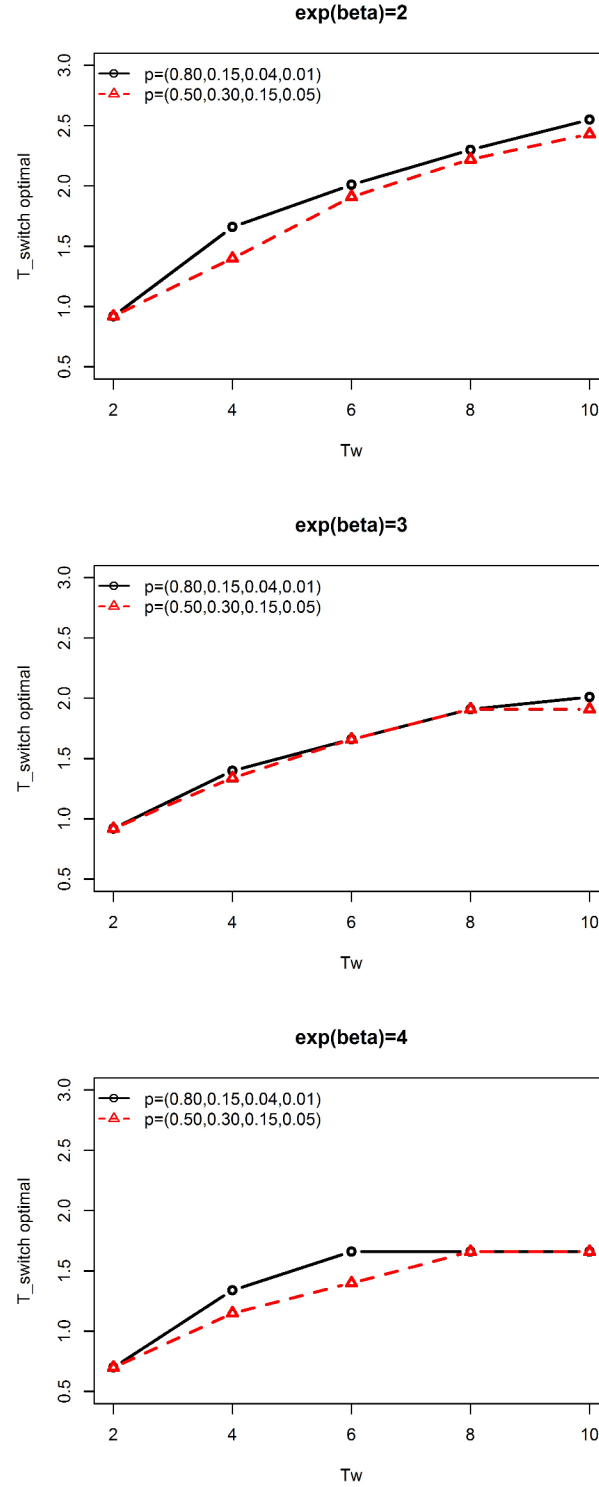


FIGURE 5.6 – Fragilité de distribution catégorielle :  $T_{\text{switch}}^0$  en fonction de  $T_w$  et de  $P$  pour  $e^\beta = 2, 3$  ou  $4$ .

le temps de changement optimal et la valeur de  $\beta$ , tandis que la durée de l'essai  $T_w$  agit positivement sur  $T_{switch}^0$ . Il faut cependant noter que le rapport entre  $T_{switch}^0$  et  $T_w$  n'est pas stable et diminue avec la valeur de  $T_w$ . Sous leurs hypothèses de travail, Buyze et Goetghebeur (2011) trouvent un temps de changement optimal correspondant à la médiane de survie. Dans ce travail, la survie n'a pas directement été étudiée comme facteur du temps de changement  $T_{switch}^0$  mais les observations faites sur l'absence d'effet significatif des différents paramètres du risque de survenue de l'évènement, traduisent clairement le fait que la survie ne peut pas être considérée comme facteur influant sur  $T_{switch}^0$ . Par ailleurs les essais tels que les essais de prévention à l'infection VIH connaissent une très forte proportion de censure, et la médiane de survie n'est jamais atteinte. A titre informatif, la proportion de censure à la fin du suivi pour  $\exp(\beta) = 2$  est autour de 50% et 55% pour les fragilités gamma et log-normale, respectivement, variant légèrement selon la variance de la fragilité. Dans les cas catégoriels étudiés, la proportion de censure en fin de suivi avoisine les 90% pour  $P = (0.80, 0.15, 0.04, 0.01)$  et 70% pour  $P = (0.50, 0.50, 0.15, 0.05)$ . Même dans les populations à haut risque d'infection, les essais de prévention VIH connaissent souvent plus de 95% de censure. Cependant, outre le problème de détermination de  $T_{switch}^0$ , le choix d'un suivi n'étant pas divisé en deux périodes égales diffère de la norme en termes de cross-over et peut alors poser un problème éthique.

Aussi, bien que l'utilisation du cross-over permette de réduire considérablement le biais, il est important de souligner qu'avec le modèle semi-paramétrique choisi, le biais zéro n'est jamais atteint. Cela pourrait s'expliquer par le fait que les sujets ayant présenté l'évènement avant le temps  $T_{switch}$  ne subissent pas le changement de groupe et créent un biais inévitable sur l'estimation de l'effet traitement.

Bien que la censure soit supposée fixe pour tous les sujets dans ce chapitre, il est possible de lui attribuer une distribution particulière. Toutefois, le changement de groupe

à un temps fixée ne permet pas une totale liberté concernant sa distribution.

# Conclusion et perspectives

En conclusion, le travail réalisé au cours de cette thèse nous a permis dans un premier temps de quantifier le biais sur l'estimation de l'effet traitement, engendré par la non-prise en compte dans l'analyse de facteurs de risques importants représentés par une variable fragilité non observée. Ainsi, selon la distribution suivie par cette fragilité, il a été observé que l'effet traitement pouvait être sous-estimé de plus de 30%, montrant ainsi l'importance de considérer la possible hétérogénéité individuelle présente dans la population de l'essai. Il s'est avéré que le biais était fortement dépendant du type de distribution de la fragilité. D'autre part, au cours de nos investigations, la forte influence de la durée de l'étude sur le biais, un facteur non étudié dans la littérature, est apparue, présentant plus particulièrement une association positive avec la valeur du biais. Ce résultat est expliqué par une interprétation de l'EMV de l'effet traitement comme étant une moyenne pondérée des rapports des risques marginaux au cours du temps. Cela nous encourage à porter une attention particulière sur les essais de longue durée, pour lesquels l'erreur peut être élevée.

En se plaçant dans le cadre des essais de prévention à l'infection VIH, le cas d'une fragilité traduisant un statut « à risque » intermittent, pouvant être rencontré dans les essais de prévention VIH, a été étudié. Il s'est avéré que là encore, le biais sur l'estimation de l'effet traitement n'était pas négligeable, à moins d'un très grand nombre de variations de la valeur de l'hétérogénéité individuelle au cours du suivi.



Tous ces résultats sont en accord avec la littérature sur le fait qu'analyser un essai randomisé en considérant un modèle à risques proportionnels ne prenant pas en compte l'hétérogénéité de la population quant au risque de survenue de l'évènement, peut engendrer une forte sous-estimation de l'effet traitement et un rejet injustifié du nouveau traitement.

Le second objectif de cette thèse était alors de présenter une réponse possible à ce problème à travers l'utilisation d'un schéma d'étude en cross-over, qui avait déjà été proposée par Auvert et al. (2011) et Buyze et Goetghebeur (2011) dans le cadre des essais de prévention à l'infection VIH. Notre étude a permis de préciser les résultats obtenus dans ces travaux grâce au calcul formel du biais asymptotique. Il a donc été montré qu'un schéma en cross-over permettait une diminution du biais relatif pouvant dépasser les 90% comparativement à un schéma parallèle classique. De plus, il s'est avéré qu'il améliorait significativement la puissance du test de comparaison, avec une efficacité asymptotique pouvant être doublée dans certaines situations. On peut alors conclure que la planification d'un schéma en cross-over, quand cela est réalisable, représente une bonne alternative pour limiter l'effet de l'hétérogénéité individuelle sur l'estimation de l'effet traitement.

Comme perspectives de travail, en se plaçant dans le cadre d'une fragilité catégorielle dont l'inclusion dans le modèle d'analyse est possible, il serait intéressant de comparer les résultats obtenus pour un schéma en cross-over sous un modèle à risques proportionnels, avec ceux obtenus pour un schéma parallèle sous un modèle à fragilité.

L'inclusion d'une fragilité dans le modèle d'analyse nécessite de poser des hypothèses fortes sur sa distribution difficilement vérifiables en pratique. O'Quigley et Stare (2002) ont d'ailleurs montré qu'au niveau populationnel, les observations faites sur une population homogène avec un effet traitement diminuant avec le temps sont proches de celles obtenues pour une population hétérogène avec un effet traitement constant. Le choix

entre les deux modèles pour l'analyse est alors arbitraire ou basé sur les informations connues ne serait-ce que partiellement sur la population de l'étude. L'utilisation du cross-over pourrait permettre d'éviter ce choix. Des études méritent cependant d'être menées sur l'impact du cross-over sur l'estimation de l'effet de traitement dans le cas où les données suivraient un modèle avec un effet traitement dépendant du temps.

Concernant le temps de changement de groupe optimal, d'autres investigations sont nécessaires afin de savoir s'il est possible de le déterminer avec précision a priori. En effet, cela nécessite donc des paramètres connus ou estimables dès la planification de l'essai.



# Bibliographie

- Aalen, O. O., Borgan, O., & Gjessing, H. (2008). *Survival and Event History Analysis : A Process Point of View*. Springer edition.
- Abbring, J. H. & Van Den Berg, G. J. (2007). The unobserved heterogeneity distribution in duration analysis. *Biometrika*, 94(1), 87–99.
- Abrahamowicz, M., Du Berger, R., Krewski, D., Burnett, R., Bartlett, G., Tamblyn, R. M., & Leffondré, K. (2004). Bias due to aggregation of individual covariates in the Cox regression model. *American Journal of Epidemiology*, 160(7), 696–706.
- Auvert, B., Sitta, R., Zarca, K., Mahiane, S. G., Pretorius, C., & Lissouba, P. (2011). The effect of heterogeneity on HIV prevention trials. *Clinical Trials*, 8(2), 144–54.
- Betensky, R. a., Louis, D. N., & Cairncross, G. (2002). Influence of Unrecognized Molecular Heterogeneity on Randomized Clinical Trials. *Journal of Clinical Oncology*, 20(10), 2495–2499.
- Bretagnolle, J. & Huber-Carol, C. (1988). Effects of omitting covariates in Cox’s model for survival data. *Scandinavian Journal of Statistics*, 15(2), 125–138.
- Buyze, J. & Goetghebeur, E. (2011). Crossover studies with survival outcomes. *Statistical Methods in Medical Research*.

- Chastang, C., Byar, D., & Piantadosi, S. (1988). A quantitative study of the bias in estimating the treatment effect caused by omitting a balanced covariate in survival models. *Statistics in Medicine*, 7, 1243–1255.
- Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. Number vol. 2 in An Introduction to Probability Theory and Its Applications. Wiley.
- Gail, M., Wieand, S., & Piantadosi, S. (1984). Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika*, 71(3), 431–44.
- Gallin, J. I., Alling, D. W., Malech, H. L., Wesley, R., Koziol, D., Marciano, B., Eisenstein, E. M., Turner, M. L., DeCarlo, E. S., Starling, J. M., & Holland, S. M. (2003). Itraconazole to prevent fungal infections in chronic granulomatous disease. *New England Journal of Medicine*, 348(24), 2416–2422. PMID : 12802027.
- Hajek, J. & Sidak, Z. (1967). *Theory of Rank Tests, First Edition*. Academia, Prague and Academic Press, New York.
- Henderson, R. & Oman, P. (1999). Effect of frailty on marginal regression estimates in survival analysis. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 61(2), 367–379.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer New York, springer edition.
- IOM (Institute of Medicine) (2008). *Methodological Challenges in Biomedical HIV Prevention Trials*. Washington, D.C. : The National Academies Press.
- Keiding, N., Andersen, P. E. R. K., & Klein, J. P. (1997). The role of frailty models and

- accelerated failure time models in describing heterogeneity. *JStatistics in Medicine*, 16, 215–224.
- Lagakos, S. W. & Schoenfeld, D. A. (1984). Properties of proportional-hazards score tests under misspecified regression models. *Biometrics*, 40, 1037–48.
- Lin, D. Y. & Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, 84, 1074–78.
- Lin, N. X., Logan, S., & Henley, W. E. (2013). Bias and sensitivity analysis when estimating treatment effects from the cox model with omitted covariates. *Biometrics*, 69(4), 850–60.
- O'Quigley, J. & Strare, J. (2002). Proportional hazards models with frailties and random effects. *Statistics in Medicine*, 21, 3219–33.
- Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., & Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *British journal of cancer*, 35(1), 1–39.
- Sangha, R., Price, J., & Butts, C. a. (2010). Adjuvant therapy in non-small cell lung cancer : current and future directions. *The Oncologist*, 15(8), 862–72.
- Schmoor, C. & Schumacher, M. (1997). Effects of covariate omission and categorization when analysing randomized trials with the Cox model. *Statistics in Medicine*, 16, 225–37.
- Schumacher, M., Olschewski, M., & Schmoor, C. (1987). The impact of heterogeneity on the comparison of survival times. *Statistics in Medicine*, 6, 773–784.

- Strauss, G. M., Herndon, J. E., Maddaus, M. a., Johnstone, D. W., Johnson, E. a., Harpole, D. H., Gillenwater, H. H., Watson, D. M., Sugarbaker, D. J., Schilsky, R. L., Vokes, E. E., & Green, M. R. (2004). Randomized Clinical Trial of adjuvant chemotherapy with paclitaxel and carboplatin following resection in Stage IB Non-Small Cell Lung Cancer (NSCLC) : Report of Cancer and Leukemia Group B (CALGB) Protocol 9633. *Journal Clinical Oncology*, 22(621s), suppl; abstr 7019.
- Strauss, G. M., Herndon, J. E., Maddaus, M. a., Johnstone, D. W., Johnson, E. a., Harpole, D. H., Gillenwater, H. H., Watson, D. M., Sugarbaker, D. J., Schilsky, R. L., Vokes, E. E., & Green, M. R. (2008). Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non-small-cell lung cancer : CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 26(31), 5043–51.
- Struthers, C. & Kalbfleisch, J. (1986). Misspecified proportional hazard models. *Biometrika*, 73, 363–9.
- Sy, J. P. & Taylor, J. M. G. (2000). Estimation in a cox proportional hazards cure model. *Biometrics*, 56(1), 227–236.
- Xu, R. & O’Quigley, J. (2000). Estimating average regression effect under non-proportional hazards. *Biostatistics*, 1(4), 423–39.





# Annexes

## Annexe A

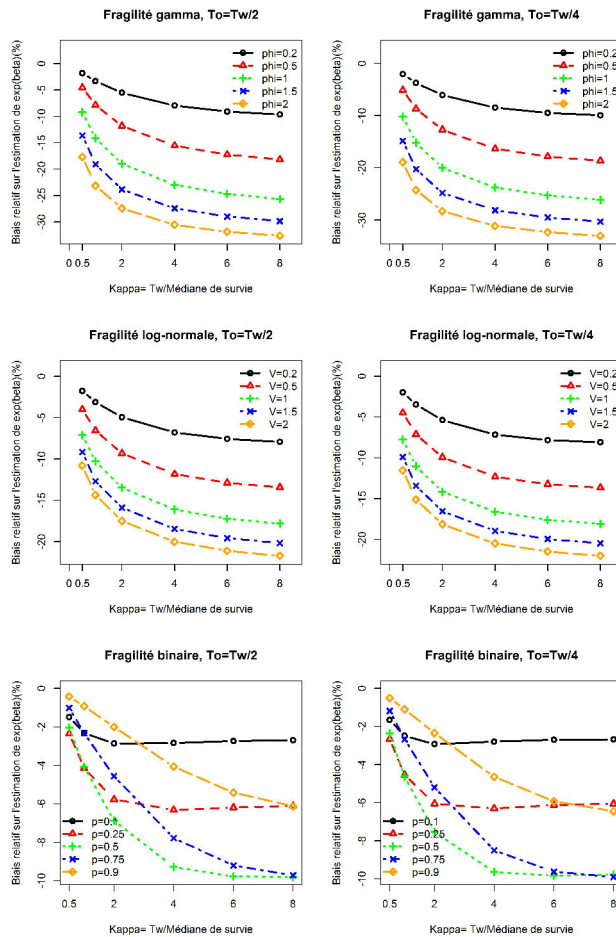


FIGURE 5.7 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité ou du paramètre  $p$ , avec  $e^\beta = 2$ . Le délai d'inclusion  $T_0 = T_w/2$  ou  $T_w/4$ .

## Annexe B

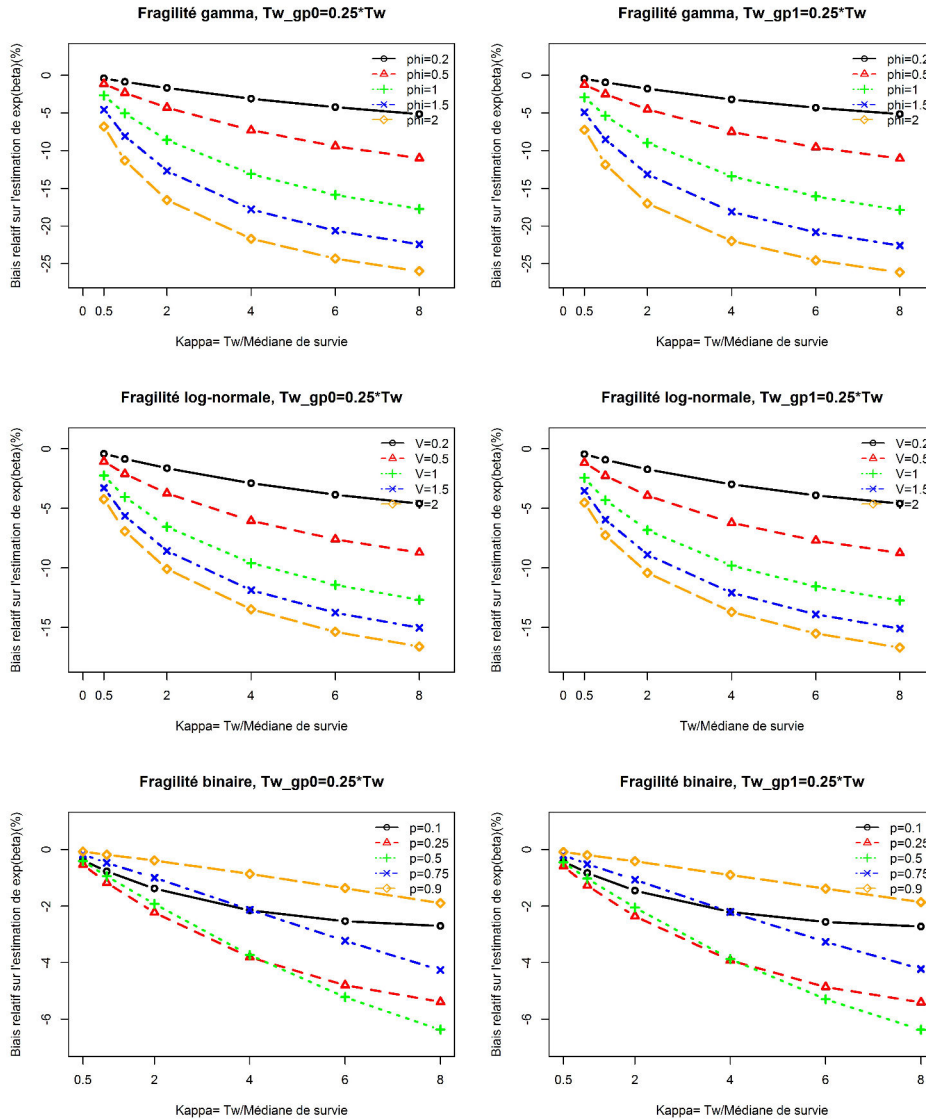


FIGURE 5.8 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et de la variance de la fragilité ou du paramètre  $p$ , avec  $e^\beta = 2$ . Délai inclusion  $T_0 = T_w/4$  dans le groupe de traitement 1 (ou 0) et  $T_0 = T_w$  dans le groupe de traitement 0 (ou 1).

## Annexe C

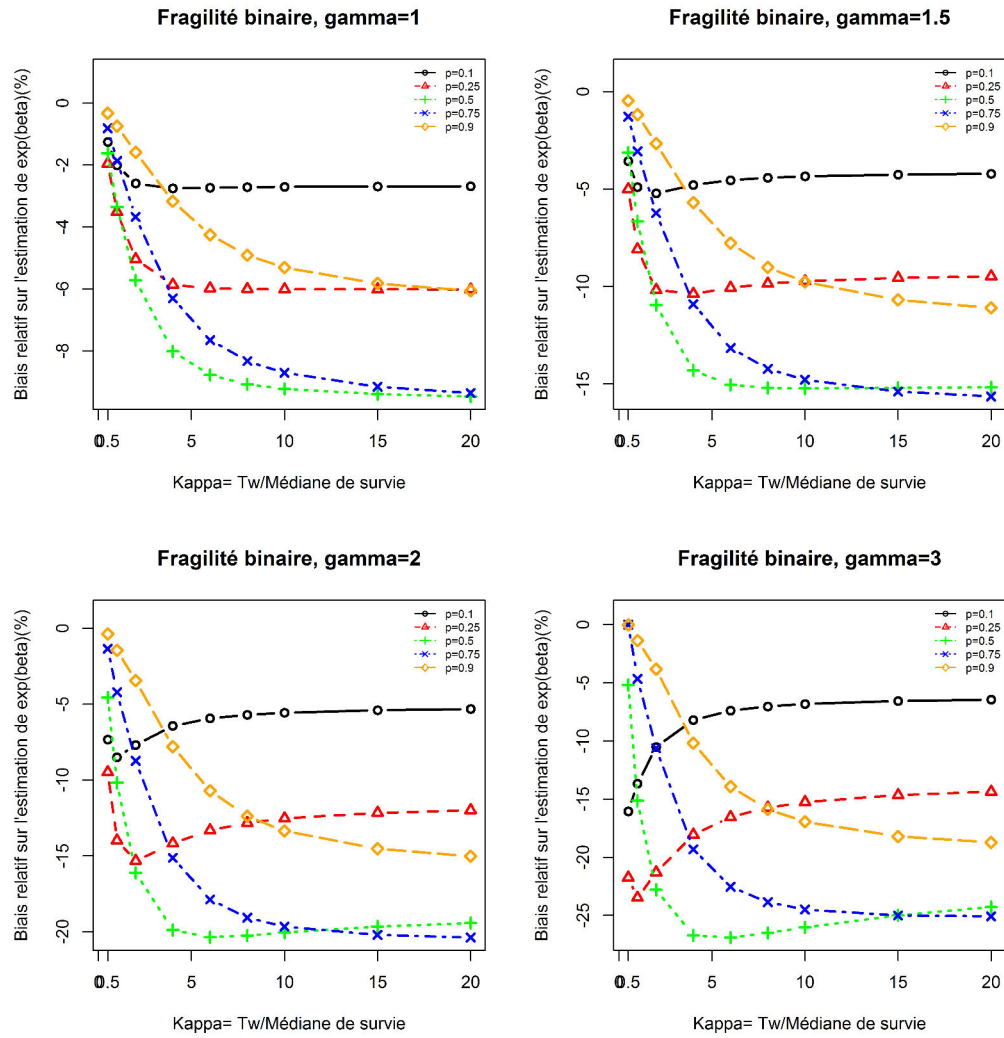


FIGURE 5.9 – Pourcentage du biais asymptotique relatif de l'estimation de  $e^\beta$  en fonction du paramètre  $\kappa$  et du paramètre  $p$ , avec  $e^\beta = 2$ . Paramètre  $\gamma = 1, 1.5, 2$  ou  $3$ .

## Annexe D : Formule de l'efficacité relative asymptotique

L'efficacité du test usuel du Logrank dans le schéma parallèle est comparée à celle du test analogue dans le schéma cross-over.

La statistique du Logrank est la statistique du score est définie par  $U(0)$  (équation (2.5)) et la condition de rejet du test pour un risque de première espèce  $\alpha$  reste :

$$\frac{U^2(0)}{\text{var}\{U(0)\}} > x_{1-\alpha}^2,$$

avec  $x_{1-\alpha}^2$  le quantile d'une loi  $\chi^2(1)$ . Par conséquent, la puissance du test est sous l'hypothèse alternative  $H_1$  définie par  $(\beta \neq 0)$  :

$$P_{H_1} \left( \frac{U^2(0)}{\text{var}\{U(0)\}} > x_{1-\alpha}^2 \right).$$

Si on note  $U^*(\beta_m)$  le membre de gauche de l'équation (2.4), la limite asymptotique de la statistique de test  $\frac{U^2(0)}{\text{var}\{U(0)\}}$  est obtenue par  $\frac{\{U^*(0)\}^2}{W(0)}$ . La puissance asymptotique du test est alors égale à :

$$P_{H_1} \left( \frac{\{U^*(0)\}^2}{W(0)} > x_{1-\alpha}^2 \right).$$

Sous l'hypothèse alternative, la statistique  $\frac{\{U^*(0)\}^2}{W(0)}$  suit une loi  $\chi^2$  décentrée dont le coefficient de non-centralité  $c$  est égal à  $E_\beta \left( \frac{\{U^*(0)\}^2}{W(0)} \right)$  qui est calculé en utilisant la vraie valeur du coefficient  $\beta$ , différent de 0. En pratique, la puissance asymptotique peut être obtenue par :

$$P_{H_1} \left( \frac{\{U^*(0)\}^2}{W(0)} > x_{1-\alpha}^2 \right) = P(Z > z_{1-\alpha} - c) + P(Z < -z_{1-\alpha} - c),$$

avec  $Z \sim N(0, 1)$  et  $z_{1-\alpha}$  le quantile d'une  $N(0, 1)$ . Le coefficient  $c$  peut être défini comme l'efficacité asymptotique du test (Hajek et Sidak, 1967). On peut alors en déduire l'efficacité relative asymptotique (ERA) du test du Logrank pour un schéma en cross-over comparativement à un schéma parallèle comme étant égale à :

$$ERA = \frac{c_{cross-over}^2}{c_{parallèle}^2},$$

avec le coefficient  $c$  calculé pour chacun des deux schémas avec les formules définies à la section 2.2.

**Title :** Survival analysis with heterogeneity between subjects in clinical trials.

**Keywords :** clinical trials, cross-over, Cox model, frailty, heterogeneity, omitted covariates, HIV prevention trials.

**Abstract :** In survival analysis, some prognostic factors can be unobserved or unavailable. In randomized clinical trials framework where two treatment groups are compared in the Cox model setting, the omission of such factors in the model leads to an under-estimation in absolute value of the treatment effect. The aim of the project is to better understand the determinants of this bias and to suggest the use of methods that could reduce it, with a particular interest in HIV prevention trials where such factors are likely to exist.

In a first step, the role of the main determinants of the bias is highlighted. While some of them have long been identified, others like trial duration have never been considered to our knowledge. The bias was calculated as the asymptotic limit of the maximum likelihood estimator of the treatment effect when the analysis is done following a proportional hazard model which no takes into account the frailty. The results show a clear increase of the bias in absolute value with the trial duration. This increase is especially marked with continuous frailty distributions, such as those which can be encountered in practice, compared to binary frailties. Also, some previous results have been confirmed.

In a second step, a frailty depending on time is considered, as it can be encountered in HIV clinical trials including “unstable” population about infection risk, as prostitutes or young people in sub-Saharan Africa. These populations present a behavioral heterogeneity linked to the change of partner whose the HIV status is unknown. The bias is estimated using a transient frailty with a simulation study reflecting real-life situations. The results show that when the frailty is regenerated during the follow-up, the bias caused by its omission, although lower than the bias obtained with a time-independent frailty, stays significant and has to be considered.

Finally, with a “stable” population whose the frailty can be supposed constant over time, the interest of the use of the “cross-over” is searched. In such trial, the subjects which have not presented the event after a given time of follow-up, change of treatment group. In HIV prevention trials framework, Auvert & al. (2011) have particularly shown a decrease of the bias with a cross-over design comparatively to a parallel design, using a categorized frailty in a simulation study. Buyze & Goetghebeur (2011) have also shown the advantages of the cross-over, particularly about the relative efficiency of the comparison test of the two treatment groups, using gamma or log-normal frailty distributions. These results are specified calculating explicitly the asymptotic bias of the hazard ratio estimate for different possible distributions of the omitted frailty, continuous or categorized. The obtained results are clearly in favor of the cross-over, with a decrease of the bias between 60% et 90% and a significant improvement of the efficiency of the comparison test. The optimal switch time and its prognostic factors are searched. It appears that it essentially depends on the trial duration and the value of the treatment effect, and is little affected by the frailty distribution.

**Titre :** Analyse de Survie en présence d'hétérogénéité entre sujets dans les essais thérapeutiques.

**Mots clés :** essais thérapeutiques, cross-over, modèle de Cox, fragilité, hétérogénéité, covariables oubliées, essais de prévention VIH.

**Résumé :** Au cours des études de survie, certains facteurs ayant un rôle pronostique peuvent être inobservés ou indisponibles. Dans le cadre des essais cliniques randomisés où deux groupes de traitement sont comparés par un modèle de Cox, « l'oubli » de tels facteurs dans le modèle entraîne une sous-estimation en valeur absolue de l'effet du traitement. L'objectif de la thèse est de préciser le rôle de différents déterminants de ce biais et de suggérer l'utilisation de méthodes pouvant le réduire, avec un intérêt particulier pour les essais de prévention à l'infection VIH où de tels facteurs peuvent exister.

L'effet des principaux facteurs pouvant influencer le biais est précisé dans une première partie de la thèse. Si certains facteurs sont connus de longue date, d'autres, comme la durée de l'essai n'ont, à notre connaissance, pas été étudiés. L'approche utilisée repose sur l'utilisation d'un « vrai » modèle de type à risques proportionnels. Dans ce modèle, l'effet des covariables « oubliées » est résumé par l'introduction dans l'expression du risque instantané d'un terme aléatoire de « fragilité » propre à chaque sujet. Le biais cherché est calculé comme la limite asymptotique, sous le modèle marginal correspondant au « vrai » modèle, du rapport des risques instantanés déduit du modèle de Cox n'incluant pas la fragilité. Les résultats montrent une nette augmentation du biais, en valeur absolue, avec la durée de l'essai. Cette augmentation est particulièrement marquée pour des distributions de fragilité continues comme celles pouvant être rencontrées en pratique, par rapport à des fragilités binaires. Par ailleurs, les résultats antérieurs de la littérature sont confirmés et précisés.

Dans un second temps, les conséquences d'une variation de la fragilité au cours du temps sur le biais sont recherchées. Plus précisément, la situation envisagée est celle qui est rencontrée dans les essais de prévention contre le VIH effectués auprès de populations « instables » vis-à-vis du risque, comme les jeunes ou les prostituées en Afrique sub-saharienne. Ces populations montrent une hétérogénéité comportementale liée aux changements de partenaires sexuels dont le statut VIH est par ailleurs inconnu. Il s'agit d'évaluer le biais en présence d'une telle fragilité « intermittente » au moyen de simulations reflétant les situations réelles. Les résultats montrent que le biais dû à l'omission de la fragilité dans l'analyse, bien qu'inférieur au biais obtenu dans le cas d'une fragilité constante au cours du temps, peut demeurer non négligeable.

Enfin, dans le cas d'une fragilité supposée constante au cours du temps, l'intérêt de l'utilisation d'un essai en « cross-over » est recherché. Dans un tel essai, les sujets exempts d'évènement après un temps de suivi donné changent de groupe de traitement. Dans le cadre des essais de prévention VIH, Auvert & al [2011] ont montré en particulier une diminution du biais avec un schéma en cross-over comparativement à un schéma parallèle classique, en utilisant une fragilité catégorielle dans une étude de simulation. Buyze & Goetghebeur [2011] ont également montré les avantages du cross-over, en particulier concernant l'efficacité relative d'un test de comparaison des deux groupes, en utilisant une fragilité de distribution gamma ou log-normale. Ces résultats sont précisés en calculant formellement le biais asymptotique dans l'estimation du risque relatif pour les différentes distributions possibles de la fragilité omise, continue ou catégorielle. Les résultats obtenus sont nettement en faveur du cross-over, avec une diminution du biais entre 60% et 90% et une amélioration de l'efficacité du test. Le temps de changement de groupe optimal est également recherché ainsi que les facteurs influant sur sa valeur. Il apparaît que celui-ci dépend essentiellement de la durée de l'essai, et est très peu affecté par la distribution de la fragilité ou la valeur de l'effet traitement.